Designing Information Provision Experiments*

Ingar Haaland Christopher Roth Johannes Wohlfart

April 29, 2021

Abstract

Information provision experiments allow researchers to test economic theories and answer policy-relevant questions by varying the information set available to respondents. We survey the emerging literature using information provision experiments in economics and discuss applications in macroeconomics, finance, political economy, public economics, labor economics, and health economics. We also discuss design considerations and provide best-practice recommendations on how to (i) measure beliefs, (ii) design the information intervention, (iii) measure belief updating, (iv) deal with potential confounds, such as experimenter demand effects, and (v) recruit respondents using online panels. We finally discuss typical effect sizes and provide sample size recommendations.

Keywords: Experimental Design, Beliefs, Information, Surveys, Obfuscation
JEL Classification: C90, D83, D91

*Ingar Haaland, University of Bergen and CESifo, email: Ingar.Haaland@uib.no; Christopher Roth, University of Warwick, briq, CESifo, CEPR, CAGE Warwick, email: Christopher.Roth@warwick.ac.uk; Johannes Wohlfart, Department of Economics and CEBI, University of Copenhagen, CESifo, Danish Finance Institute, e-mail: johannes.wohlfart@econ.ku.dk; The activities of the Center for Economic Behavior and Inequality (CEBI) are funded by the Danish National Research Foundation. We thank the editor (Steven Durlauf) and three anonymous referees for extremely useful suggestions. We also thank Peter Andre, Yan Chen, Guillermo Cruces, Adeline Delavande, Philipp Doerrenberg, Lukas Hensel, Mitchell Hoffman, Philipp Lergetporer, Michel Marechal, David McKenzie, Brendan Nyhan, Ricardo Perez-Truglia, Peter Schwardmann, Gaurav Sood, Diego Ubfal, Pierre-Luc Vautrey, and Basit Zafar for very useful suggestions. Cecilie Christensen, Hrishikesh Iyengar, and Florian Schoner provided excellent research assistance.

1 Motivation

Standard economic theories usually understand choices as a combination of three factors: preferences, constraints, and beliefs. Information enters the choice environment indirectly by affecting beliefs and (perceived) constraints. For instance, in the context of a firm owner choosing how much to invest, information about historical returns could affect beliefs about the return on investment while information about loan opportunities could affect borrowing constraints. The goal of economic experiments is typically to change some features of the choice environment to study how choices are made. Information experiments achieve this by varying the information set available to economic agents.

Information provision experiments allow researchers to provide cleanly identified evidence by only varying one feature of the information set. This in turn provides researchers with tools to test either basic assumptions of models or differentiate between different theoretical mechanisms. For example, Bursztyn et al. (2014) use information treatments to study the relative importance of “social learning” and “social utility” in the context of peer effects in financial decisions, while Allcott and Taubinsky (2015) use information treatments to identify the importance of different behavioral biases in the lightbulb market.

One powerful possibility opened up by information experiments is to generate exogenous variation in perceptions of real-world environments, which allows answering policy-relevant questions. Information provision experiments allow changing perceptions of real-world phenomena which themselves cannot be directly changed. For instance, in labor economics, an important policy question is to what extent students internalize market returns to education when making educational choices. While researchers cannot manipulate the returns to education, they can provide information to generate exogenous variation in the perceived returns to education (Jensen, 2010). When studying attitudes towards immigration, it is impossible to change the characteristics of immigrants, but researchers can vary perceptions of the immigrant population by correcting people's misperceptions (Grigorieff et al., 2020). Similarly, researchers cannot manipulate intergenerational mobility or influence the state of the macroeconomy, but it is possible to change perceptions of intergenerational mobility (Alesina et al., 2018c) or the perceived likelihood of a recession (Roth and Wohlfart, 2020). Finally, researchers cannot manipulate social norms, but information provision experiments can be used to study the causal effect of perceived social norms on behavior (Bursztyn et al., 2020b).

The opportunities provided by information experiments to test economic theories and answer policy-relevant questions have made them popular in economics over the last few years. As shown in Figure 1, the number of information provision experiments published in leading economics journals has strongly increased over the last ten years. This growth demonstrates the increasing importance of information provision in experimental work. In this article, we review the growing literature on information experiments in economics with a particular focus on methodological questions. Many of the methodological questions discussed in this review also extend beyond information provision experiments, and have relevance for the design of other types of experiments or for collecting non-experimental survey data.

The paper proceeds as follows: Section 2 summarizes areas in which information experiments have been widely applied. Section 3 outlines best-practice recommendations for the measurement of beliefs. Section 4 discusses the design of the information intervention.

Section 5 outlines important aspects of the measurement of belief updating. Section 6 discusses best-practice recommendations for mitigating concerns about experimenter demand effects. Section 7 discusses online samples that are commonly used for information provision experiments. Section 8 discusses typical effect sizes and provides sample size recommendations. Finally, Section 9 offers concluding remarks.

2 Major applications

In this section, we provide an overview of areas in economics in which information provision experiments have been widely applied. This review is necessarily incomplete, and focuses on applications in public economics, political economy, macroeconomics, household finance, and labor, education and health economics.[1]

Public Economics Information provision experiments are used in many areas of public economics. Chetty and Saez (2013) conduct an experiment with 43,000 Earned Income Tax Credit (EITC) recipients, in which a random subset received personalized information about the EITC schedule. Chen et al. (2010) find that personalized social information can significantly increase digital public goods provision, in the form of user moving ratings and community database maintenance. Chen et al. (2017) conduct a field experiment with 75,000 drivers, establishing that social information can reduce traffic violations. Doerrenberg and Peichl (2018) examine how social norms affect tax morale, and Blesse et

[1] Our review does not include information provision experiments operating in a laboratory setting in which respondents receive information about features of the laboratory environment or the behavior of other participants in the lab. The review also does not feature work studying the role of the media in shaping beliefs and behavior.

al. (2019) study how beliefs shape preferences for tax simplification.

Bérgolo et al. (2017) and Doerrenberg and Schmitz (2015) examine how firms respond to information about audit probabilities, and Bott et al. (2020) study whether people's tendency to evade taxes responds to information about detection probability and moral appeals. Similarly, Perez-Truglia and Troiano (2018) examine how information on financial penalties, shaming penalties, and peer comparisons shape tax delinquents' future repayment rates. De Neve et al. (2021) study the impact of deterrence, tax morale, and simplifying information on tax compliance. Jessoe and Rapson (2014) show that information about residential electricity usage makes households more responsive to temporary price increases. Finally, a literature in behavioral public economics has studied how misperceptions regarding the fuel economy affect consumers' purchasing decisions (Allcott, 2013; Allcott and Knittel, 2019; Allcott and Taubinsky, 2015).

Political Economy Information experiments are also commonly used to study how information affects policy attitudes, such as people's demand for redistribution (Alesina et al., 2018c; Chen et al., 2016; Cruces et al., 2013; Fehr et al., 2021, 2019; Gärtner et al., 2019; Hoy and Mager, 2018; Karadja et al., 2017; Kuziemko et al., 2015), their support for government spending (Lergetporer et al., 2018a; Roth et al., 2021a), their views on educational inequality (Lergetporer et al., 2020) and tuition fees (Lergetporer et al., 2016), their support for immigration (Alesina et al., 2018a; Bansak et al., 2016; Barrera et al., 2020; Facchini et al., 2016; Grigorieff et al., 2020; Haaland and Roth, 2020; Hopkins et al., 2019; Lergetporer et al., 2017), their tendency to discriminate against immigrants (Alesina et al., 2018b), their support for affirmative action (Haaland and Roth, 2021; Settele, 2020), or affective party polarization (Ahler and Sood, 2018). In the context of the coronavirus pandemic, Settele and Shupe (2020) study the role of beliefs for supporting lockdown measures and Rafkin et al. (2021) study determinants of inference from official government projections.

Information experiments are also conducted to better understand the demand for news and the implications of media exposure for behavior. Chopra et al. (2021) study how perceived informativeness affects people's demand for economic and political news. Bursztyn et al. (2020e) study how the common knowledge of rationales (which are usually supplied by the media) affects the public expression of xenophobia through the lens of a signaling model.

In the context of natural field experiments, researchers have used information treatments to study voting behavior (Aker et al., 2017; Cruz et al., 2021, 2018; Gerber et al., 2020; Kendall et al., 2015; Orkin, 2019) or to study strategic behavior of political activists (Hager et al., 2020, 2021) and protesters (Cantoni et al., 2019; Hager et al., 2019).

Experiments studying the effects of information campaigns in the context of political behavior have also been widely applied in developing country settings. Armand et al. (2020) test whether information can counteract the political resource curse in Mozambique. Acemoglu et al. (2020) study whether information about improved public services can help build trust in state institutions and move people away from non-state actors in Pakistan. Khan et al. (2020) document that information about past state effectiveness has a limited impact on support for policy, perceptions of state capacity, and trust in the state in Pakistan. Finally, Banerjee et al. (2018) show that mailing cards with program information to targeted beneficiaries reduces "leakage" in redistribution programs due to local officials not implementing government programs.

Macroeconomics Information provision experiments have been widely used in macroeconomics to study how households and firms form expectations about inflation (Armantier et al., 2016; Binder and Rodrigue, 2018; Cavallo et al., 2017; Coibion et al., 2021b, 2018, 2020d,e), house prices (Armona et al., 2019; Fuster et al., 2020), interest rates (Coibion et al., 2020a; Link et al., 2021) or the broader economy (Coibion et al., 2020c). Another set of studies have applied information provision experiments to examine the causal effect of macroeconomic expectations on behavior. For instance, these papers have studied how households' spending decisions are affected by expectations about GDP growth (Coibion et al., 2021a; Roth and Wohlfart, 2020), house prices (Qian, 2019), or inflation (Coibion et al., 2019a), and how inflation expectations influence firms' decisions (Coibion et al., 2019c). In a developing country context, Galashin et al. (2020) examine how information about inflation or the exchange rate affects consumer spending as measured in administrative credit card data. Roth et al. (2021b) study how households' perceptions of their exposure to macroeconomic risk causally affect information acquisition. Finally, information experiments have been used to study the effectiveness of different forms of policy communication (Coibion et al., 2019b; D'Acunto et al., 2020). In the context of the Covid-19 pandemic, Coibion et al. (2020b) and Binder (2020) study how provision of information about policy responses shapes households' macroeconomic expectations and spending plans.

Household Finance Research in household finance has studied the effects of information provision on retirement savings (Beshears et al., 2015; Dolls et al., 2018). Moreover, Bursztyn et al. (2019) examine how moral appeals affect debt repayment. Bursztyn et al. (2014) study the mechanisms underlying peer effects in financial decisions. Bottan and Perez-Truglia (2020a) study the causal effect of home price expectations on the timing of home sales using a large-scale field experiment featuring administrative data. Laudenbach et al. (2021) use an information experiment to study the causal effect of subjective beliefs about stock returns on investment choices measured in administrative account data. In the context of the coronavirus pandemic, Hanspal et al. (2020) provide experimental evidence that beliefs about the duration of the stock market recovery shape households' expectations about their own wealth and their planned investment decisions and labor market activity.

Labor and education economics Information provision experiments have been applied to study job search (Abebe et al., 2020; Altmann et al., 2018; Belot et al., 2018, 2019; Carranza et al., 2020; Franklin, 2017), social norms (Bursztyn et al., 2020b), educational aspirations (Lergetporer et al., 2018b), schooling decisions (Jensen, 2010), major choice (Bleemer and Zafar, 2018; Conlon, 2019; Wiswall and Zafar, 2014), postgraduate enrolment (Berkes et al., 2019) as well as school choice (Ajayi et al., 2017; Andrabi et al., 2017). Researchers have shown that information about school quality affects parental investment decisions (Greaves et al., 2019) and that parents' beliefs about children's ability affect their educational investments (Dizon-Ross, 2019). Coffman et al. (2017) highlight that information about peers' choices can affect job choice. Researchers in behavioral labor economics have also studied how information provision about peers affects people's work morale and labor market behavior (Card et al., 2012; Cullen and Perez-Truglia, 2018). In agricultural economics, information provision experiments are also widely applied; for example, Hanna et al. (2014) study the effects of information on farmers' behavior. In the context of migration, Baseler (2021) studies how perceived returns to migration shape migration decisions in Kenya, while Shrestha (2020) studies how information about the potential risks of dying and potential wages from working abroad affects actual migration decisions in Nepal. Humphries et al. (2020) study the role of information frictions for access to the Paycheck Protection Program in the context of the coronavirus pandemic.

Health economics Information provision experiments have also been widely used to study information relevant for health behaviors. In the context of the US, Nyhan and Reifler (2015) and Nyhan et al. (2014) study the effects of information campaigns about vaccines. Alsan and Eichmeyer (2021) study persuasion regarding the medical benefits of influenza vaccination with a particular focus on racial identity. Barari et al. (2020) study public health messaging and social distancing in the context of the coronavirus pandemic, while Fetzer et al. (2020) and Akesson et al. (2020) study perceptions of the pandemic risk factors. Faia et al. (2021) use an information experiment to study biases in information selection and processing in the context of the pandemic.

In developing country settings, Kremer and Miguel (2007) document muted effects of information on how to avoid worm infections.[2] Fitzsimons et al. (2016) find that information provision to mothers in Malawi increases children's food consumption. Carneiro et al. (2020) study an intervention targeting early life nutrition, which also provides nutritional information. Banerjee et al. (2015) examine how information affects take-up of Double Fortified Salt. Bennear et al. (2013) examine how household drinking-water choices were

[2]For a review of this literature in developing country settings, see Dupas and Miguel (2017).

affected by two different messages about risk from naturally occurring groundwater arsenic. Madajewicz et al. (2007) show that information about well safety regarding arsenic levels results in large switches from unsafe to safe water sources. Levine and Kinder (2004) discuss the success of on oral rehydration information campaign for diarrhea treatment.

A substantial amount of research has been conducted on the role of beliefs in sexual behaviors. Dupas (2011) and Kerwin (2018) examine how information about HIV risks affects sexual behaviors. Ciancio et al. (2020) study how mortality risk information changes survival expectations and sexual behaviors in Malawi. Chong et al. (2013) study the impact of an online sexual health education course provided through schools in Colombia. Jamison et al. (2013) examine how information about sexual and reproductive health affects knowledge and sexual behavior in a general population sample in Uganda. Shacham et al. (2014) study how information about the relationship between circumcision and HIV status affects sexual behavior in Malawi. Miller et al. (2020) find that simply informing women about pregnancy risk increases stated intentions to use contraception substantially. Researchers has also studies how learning about one's HIV status affects sexual behavior (Delavande and Kohler, 2012; Gong, 2015).

As Kremer et al. (2019) note, most of studies in the context of developing countries focus on the effects of information on behavior and not beliefs. We believe that the elicitation of beliefs would be helpful to more clearly understand why certain information interventions are more successful in changing behavior than others. This is especially important in the context of health behaviors, where trust in information may vary substantially depending on prior belief and the identity of the sender.

3 Measuring Beliefs

Information provision experiments aim to study the effect of information on people's beliefs or to generate exogenous variation in beliefs to study the effect of beliefs on other outcomes. This section discusses whether one should measure prior beliefs before the information provision and posterior beliefs after the information provision. It also discusses issues related to the measurement of beliefs, including advantages and disadvantages of measuring qualitative or quantitative point beliefs versus probabilistic beliefs, the use of external benchmarks for the elicitation of beliefs, the framing of belief elicitations, and techniques on how to deal with measurement error. Finally, we review the measurement of beliefs using hypothetical vignettes.

3.1 Eliciting prior and posterior beliefs?

There are several advantages to eliciting prior beliefs in information provision experiments. First, eliciting prior beliefs about the provided piece of information allows researchers to estimate heterogeneous treatment effects by prior beliefs. This is particularly relevant in designs with a pure control group (that is, a control group that does not receive any information). Depending on their priors, groups of participants may update their beliefs in different directions in response to the information, leading to a muted average treatment effect. For instance, consider the experiment by Cruces et al. (2013), which gives people information about their relative position in the income distribution. Since some people overestimate their position while other people underestimate their position, providing accurate information will lead overestimators and underestimators to update their beliefs in opposite directions. Eliciting prior beliefs is, therefore, necessary to make a directional prediction about how different groups should update their beliefs and change their behavior in response to the information. Furthermore, even if all respondents update their beliefs in the same direction, analyzing heterogeneity by prior beliefs allows the researcher to assess whether treatment effects on the outcome of interest are larger for respondents who received a larger information shock. Such heterogeneous effects by prior beliefs are often viewed as evidence that treatment effects are driven by genuine changes in beliefs rather than priming (see Section 4.5). Second, when the outcome of interest is an expectation, eliciting prior beliefs related to this outcome allows the researcher to estimate learning rates from the information (see Section 8). Third, eliciting either type of prior beliefs increases statistical power for detecting treatment effects (Clifford et al., 2020).

Eliciting posterior beliefs is important in settings where there is a direct interest in studying the effect of information on these beliefs. Moreover, the measurement of posterior beliefs is necessary to learn about the size of the first stage in settings where information provision experiments are used to study the causal effects of beliefs on other outcomes. In settings where respondents are provided with information about facts (e.g., Roth et al., 2021a), eliciting posteriors primarily serves to measure trust in or attention to the information. As such, eliciting posteriors is less strictly needed than in designs where respondents receive a potentially noisy signal about a variable (e.g., Roth and Wohlfart, 2020), where posteriors are used to assess how informative respondents find the provided signal.

A potential downside of designs measuring both priors and posteriors about the same object is that such within-designs potentially lead to stronger experimenter demand effects (see Section 6). Alternatively, respondents may be subject to consistency bias in their survey responses (Falk and Zimmermann, 2012), leading to a muted effect of information in within-designs. However, Roth and Wohlfart (2020) do not find any significant effect of eliciting priors on the estimated average learning rate in the context of information about macroeconomic risk. Similarly, Clifford et al. (2020) find little evidence of bias in estimated treatment effects due to the elicitation of prior beliefs in survey experiments on political attitudes. Moreover, in designs with a pure control group, being asked the same question twice might confuse respondents in the control group. One remedy is to use a different elicitation mode for the posteriors compared to the priors (Coibion et al., 2019b) or to elicit post-treatment beliefs about a related but different outcome (Haaland and Roth, 2021).

3.2 Qualitative, quantitative or probabilistic beliefs?

How exactly should one measure beliefs? Should one measure beliefs using qualitative or quantitative survey questions? Should one measure point estimates on quantities or probabilistic beliefs in which people attach probabilities to different states of the world occurring?[3]

Qualitative beliefs: verbal response scales One way to measure beliefs is to present respondents with verbal response scales, e.g. reaching from "very low" to "very high," or from "strongly agree" to "strongly disagree." Such belief measures have the simple advantage that the response options are easy to understand for respondents, but a clear disadvantage is that they are not easily interpersonally comparable, which can result

[3]See Delavande et al. (2011) and Delavande (2014) for excellent overviews on the measurement of subjective beliefs with a particular focus on developing country settings.

in severe identification challenges (Bond and Lang, 2019). For instance, in the context of measuring beliefs about the size of the immigration population, people might hold systematically different views on whether a given fraction of immigrants in the population is “very low” or “very high.” Such differences in the interpretation of qualitative response options can be driven by partisanship, as shown by Gaines et al. (2007) for the case of beliefs about the Iraq war. Moreover, verbal response scales are relatively crude and therefore limit the extent of information that can be conveyed (Manski, 2018). Furthermore, with qualitative beliefs, it is often theoretically ambiguous in which direction people should update their beliefs in response to an information treatment. For instance, to manipulate perceptions about the size of the immigration population in the United States, one could inform treated respondents that 12 percent of the US population are immigrants (Grigorieff et al., 2020; Hopkins et al., 2019). Without a quantitative pre-treatment beliefs measure, it is not clear whether treated respondents should revise their beliefs about the size of the immigration population upwards or downwards in response to this information. At the same time, given their advantages in terms of simplicity, including qualitative belief measures can at least serve as a robustness check for results based on quantitative belief data.

Qualitative beliefs: open-ended questions It is also possible to use open-ended questions to measure beliefs (Andre et al., 2019; Bursztyn et al., 2020e; Stantcheva, 2020). The key advantage of such open-ended questions is that respondents are not primed by the available response options. In other words, open-ended questions enable researchers to directly measure what “comes to mind.” They therefore allow researchers to shed light on people's attention allocation. For example, Andre et al. (2019) study which propagation mechanisms come to households' and experts' minds when thinking about canonical macroeconomic shocks. In the context of macroeconomic expectation formation, Leiser and Drori (2005) and D'Acunto et al. (2019) study people's associations with inflation using open-ended text questions. Stantcheva (2020) examines what considerations people have in mind when thinking about a given policy. Bursztyn et al. (2020e) use such an open-ended elicitation to study inference about the motives for xenophobic expression. Using a pre-registered text analysis procedure and hand-coding of the qualitative responses, they use this data for studying inference. They validate their open-ended question with a structured belief measure and establish strong correlations. In the context of information provision experiments, open-ended questions have two main purposes. First, they can be used as a validation check for the quantitative belief data (as in Bursztyn et al. 2020e). Second, they can be used to study whether providing information changes people's attention to a given topic.

Quantitative point beliefs Quantitative point beliefs, where respondents are asked to state their beliefs on a numerical scale, have the advantages of interpersonal comparability while still being relatively straightforward for respondents to understand, but they do not allow for individuals to express their uncertainty about outcomes. It is therefore a good practice to add a second qualitative question on how sure or confident people were in their previous answer. For instance, such qualitative measures of confidence can be used for tests of Bayesian updating (Armona et al., 2019; Roth and Wohlfart, 2020)[4] or to examine whether

[4]Since confidence in priors is not randomly assigned and is likely correlated with other variables, such tests are often only suggestive.

subjective measures of confidence are related to the accuracy of people's beliefs (Graham, 2018). Furthermore, eliciting confidence allows the researcher to differentiate between strong misperceptions and a lack of knowledge. A second disadvantage of point beliefs is that it is unclear which feature of their subjective belief distribution over potential future outcomes respondents report. While researchers often implicitly or explicitly interpret point beliefs as the mean over the respondent's subjective distribution, respondents may report their median or mode belief.[5] Lastly, people's point beliefs might be sensitive to question framing (Eriksson and Simpson, 2012).

Probabilistic beliefs In probabilistic belief elicitation, respondents state probabilities for the occurrence of different mutually exclusive events. Such probabilistic elicitations have the advantage that there is a well-defined absolute numerical scale that is interpersonally and intrapersonally comparable (Manski, 2018). Probabilistic elicitations were pioneered by Manski (2004) and have been widely and successfully applied in some areas in economics, such as labor economics, and the economics of education (Attanasio et al., 2021; Boneva and Rauh, 2017; Boneva et al., 2019; Wiswall and Zafar, 2016, 2017) as well as health economics (Delavande and Kohler, 2009). These measures allow researchers to directly compute a measure of uncertainty as well as the mode, the median and the mean. Probabilistic data also enable researchers to characterize the nature of updating beyond showing average learning rates, or to document biases in updating. For instance, Barron (2021) uses a probabilistic elicitation to test whether people update their beliefs in the financial

[5]For instance, De Bruin et al. (2011) show that survey respondents' point forecasts about future inflation or future wage growth are not consistently associated with means constructed from individual-level subjective probability distributions over future inflation or wage growth, but are often associated with the median or other measures of central tendency of respondents' reported distribution.

domain in line with a Bayesian benchmark. Directly measuring people's uncertainty has recently received additional attention in the context of abstract choice and updating tasks as well as survey expectations. Enke and Graeber (2019) propose a measurement of cognitive uncertainty and show people who are cognitively uncertain implicitly compress probabilities towards a cognitive default of 50:50 in binary state spaces.

One drawback of probabilistic scales is that a large fraction of the population has difficulties in understanding and interpreting probabilities (Tversky and Kahneman, 1974). A second drawback is that people's stated beliefs are typically influenced by how the outcomes are categorized (Benjamin et al., 2017). A third drawback is that probabilistic questions are more time-consuming and taxing for respondents, which makes the experiment longer and potentially induces higher attrition or a higher fraction of missing responses. Some survey providers might also object to the use of probabilistic questions as they might confuse respondents. However, best-practice recommendations, such as the use of visual interfaces, have been shown to mitigate some of these problems, even when working with low-literacy populations (Delavande, 2014; Delavande et al., 2011).

3.3 Benchmarks

One approach measures beliefs about objects of interest for which there is an objective external benchmark. For instance, in the context of income inequality, one can elicit beliefs about the income share going to the top 1 percent income earners rather than a general question about whether income inequality is "high" or "low." Measuring subjective beliefs about quantities with well-defined benchmarks has several advantages. First, by eliciting beliefs about a well-defined benchmark the experimenter fixes beliefs about the environment and imposes additional structure on the responses. This in turn may lower heterogeneity in how the question is interpreted and thereby reduce measurement error and make responses across participants more comparable. Second, it allows one to characterize the extent of biases in beliefs compared to the benchmark. Third, it enables one to incentivize the belief elicitations in a transparent way. Fourth, the availability of benchmarks allows for the provision of information treatments that are tightly connected with the belief elicitation. Recent applications of belief elicitations reliant on benchmarks are studies on social norm elicitations (Krupka and Weber, 2013), racial discrimination (Haaland and Roth, 2021), intergenerational mobility (Alesina et al., 2018c), immigration (Alesina et al., 2018a; Grigorieff et al., 2020; Haaland and Roth, 2020), or infectious disease spread (Akesson et al., 2020; Fetzer et al., 2020).

3.4 Framing of belief elicitations

In settings in which respondents are relatively experienced they are capable of accurately assessing economic quantities. For example, respondents are relatively good at assessing the price of gas (Ansolabehere et al., 2013). However, in settings in which respondents are relatively unfamiliar, there will be higher levels of measurement error especially when respondents are unsure about the response scale, for example in the context of unemployment estimates. However, careful framing of questions can reduce measurement error. For example, the provision of anchors which convey information about the response scale can reduce measurement error (Ansolabehere et al., 2013). For instance, Roth et al.

(2021a) measure beliefs about the debt-to-GDP ratio in the US using different historical or cross-country anchors, and show that the provision of an anchor reduces the dispersion of beliefs and rounding.

3.5 Multiple measurement

Many belief elicitations involve asking difficult questions to respondents. The cognitive strain in turn may induce measurement error. How can researchers mitigate the extent of measurement error? Gillen et al. (2019) propose an IV approach, which leverages multiple measurements to deal with classical measurement error. We believe that this is particularly important in the context of (quantitative) belief measurement. When reducing classical measurement error is important, researchers ideally should measure their belief of interest using (i) a qualitative survey question, (ii) a quantitative point estimate, and (iii) a probabilistic question in which respondents attach probabilities to mutually exclusive states of the world. These multiple measurements in turn can be leveraged to employ the IV methods that help to deal with measurement error (Gillen et al., 2019). For instance, Giglio et al. (2021) apply such an IV approach in the context of survey expectations about stock returns, using both point beliefs and subjective probability distributions. However, since multiple measurements might be cognitively taxing for respondents, their benefits must be weighed against the risk of increasing survey fatigue or higher attrition rates. Moreover, this approach cannot be used in the case of non-classical measurement error. Finally, randomly assigned information treatments can also be used to instrument beliefs, and thereby deal with measurement error to some extent (for a discussion of such an IV approach, see Section 8).

3.6 Incentives

When eliciting beliefs with an objective external benchmark, it is possible to provide accuracy incentives to encourage higher effort and more truthful responses. For instance, with a discrete outcome, one can promise the respondents a monetary reward if they answer the question correctly. Similarly, with a continuous outcome, one can offer the respondents a monetary reward if their answer is within some percentage range of the correct answer. The advantages of these mechanisms are that they are simple to explain to respondents and provide stark incentives to provide correct answers. The disadvantage of these mechanisms is that they are only incentive-compatible for eliciting the mode of a respondent's belief distribution. While respondents in some situations might be perfectly willing to provide truthful responses even in the absence of monetary incentives, incentives might be especially important in political settings where respondents might form motivated beliefs or receive expressive utility from stating beliefs that are consistent with their partisan leanings (e.g., Democrats stating a low unemployment rate under a Democratic president). Consistent with survey respondents stating beliefs in a motivated way, incentives have been shown to reduce partisan bias in people's stated beliefs about economics and politics (Bullock et al., 2015; Prior et al., 2015). For example, the partisan gap in beliefs about the current unemployment rate shrinks when respondents receive prediction incentives.[6] Relatedly, Settele (2020) shows that gender differences in reported

[6]This evidence suggests that prediction incentives can lower motivated errors, thereby reducing non-classical measurement error.

beliefs about the gender wage gap shrink in the presence of incentives. Peterson and Iyengar (2020) find a moderate reduction in partisan differences in beliefs on topics such as climate change, immigration, or firearms when survey respondents are provided with incentives and Berinsky (2018) finds small effects of incentives on respondents' tendency to endorse political rumors. Allcott et al. (2020), on the other hand, find no effect of incentives on partisan differences in beliefs about the coronavirus pandemic. Trautmann and van de Kuilen (2014) find that incentives do not improve the accuracy of people's predictions about the behavior of others in lab games. In the context of macroeconomic forecasting, it has been shown that non-incentivized survey reports strongly correlate with incentivized belief measures (Armantier et al., 2015) and that incentives do not have any statistically significant effects on reported beliefs (Roth and Wohlfart, 2020). Similarly, Andre et al. (2019) find muted effects of incentives on the accuracy of macroeconomic beliefs, even though response time significantly increases. Moreover, Hoffman and Burks (2020) find no effect of incentives on workers' tendency to over-estimate their productivity. Finally, Grewenig et al. (2021) provide mixed evidence on the relevance of incentives in shaping accuracy. Their evidence highlights that incentives have similar effects as a prompt to google the statistic of interest.[7] This highlights the potential undesirable side-effects of incentives when the information of interest is publicly available.

[7] This issue might be mitigated by using software that prevents people from going outside of their current browser. Alternatively, one may ask respondents not to use outside information, but it is also conceivable that such messages might themselves have backfiring effects. In cases where it is very important to have tight control over the information environment, information provisions experiments can also be conducted in the laboratory (see e.g., Bursztyn et al. 2020b for an example of an information provision experiment performed in a laboratory setting).

Eliciting probabilistic beliefs with incentives If the goal is to elicit probabilistic beliefs in an incentive-compatible way, one needs to use a proper scoring rule. A commonly used method is the quadratic scoring rule (QSR) which provides incentive-compatibility for risk-neutral agents (Brier, 1950). Palfrey and Wang (2009) and Wang (2011) provide evidence that the QSR induces more accurate and better calibrated forecasts than improper scoring rules or non-incentivized belief elicitations, respectively. More recently, the binarized scoring rule (BSR) by Hossain and Okui (2013) has become a popular alternative to the QSR. Compared to the QSR, in which the size of the reward depends on the accuracy of people's estimates, the BSR offers a fixed reward in which the chance of winning the reward increases in the accuracy of people's estimates. This makes the BSR incentive-compatible irrespective of risk preferences. A clear disadvantage of both mechanisms is that they suffer from flat incentive structures in which relatively large deviations from truthful reporting generate only modest changes in the expected rewards. Furthermore, Danz et al. (2020) provide evidence that the complex incentive schemes could lead to less truthful reporting by making respondents misunderstand the incentive scheme. Specifically, they find a strong increase in truthful reporting in a treatment without any information about how incentives are determined compared to a baseline condition providing full information about how incentives are determined by a BSR. Based on this, we consider it best-practice to simply inform respondents that it is in their best interests to provide truthful responses when using a proper scoring rule and only provide the mathematical details to respondents who express an explicit interest, e.g. in the form of a clickable pop-up box.

Taken together, while incentives seem to be important when eliciting beliefs in the political domain, incentives seem to have little effect on stated beliefs in non-political domains.[8] Furthermore, incentives could backfire when the true answer can easily be googled or when the complexity of the incentive structure makes respondents misunderstand the payoff structure.

3.7 Hypothetical vignettes

Another approach to measuring beliefs is to ask respondents to make predictions about an outcome under different hypothetical scenarios. The use of such hypothetical vignettes is an increasingly popular approach to measure beliefs in contexts that are difficult to study in a real-world setting, such as education and human capital (Attanasio et al., 2021; Boneva and Rauh, 2017, 2018; Delavande and Zafar, 2019; Kiessling, 2021; Wiswall and Zafar, 2017), preferences over wealth taxation (Fisman et al., 2017), and beliefs about the effects of macroeconomic shocks (Andre et al., 2019). Hypothetical vignettes, in the form of conjoint experiments where many different attributes are simultaneously randomized, are widely used to study preferences over different types of immigration (Bansak et al., 2016; Hainmueller and Hopkins, 2014; Hainmueller and Hiscox, 2010). Hainmueller et al. (2015) show that the responses in the vignettes are highly predictive of real-world behaviors.

Hypothetical vignettes have the advantage of allowing the researcher more control over the context specified to respondents. Potential disadvantages of hypothetical vignettes include that the hypothetical nature may lower respondents' effort or induce experimenter

[8] However, most of this evidence is based on relatively small stakes and it is an open question whether high-stake incentives would change stated beliefs more strongly. Furthermore, there is some evidence suggesting that incentives improve the accuracy of forecasts in the context of economic games (Wang, 2011) and incentives have also been shown to increase performance in memory and recall tasks (Camerer and Hogarth, 1999).

demand effects. Furthermore, when designing hypothetical vignettes, it is important to consider whether experimentally manipulating an attribute also changes beliefs about other background characteristics (Dafoe et al., 2018). For instance, manipulating whether an immigrant is described as "motivated to find work" or "not motivated to find work" might not isolate economic concerns about immigration if the manipulation also changes beliefs about how likely they are to fit in culturally. Finally, it may be cognitively challenging for respondents to think in hypotheticals, which could in turn increase measurement error and reduce external validity.

4 Designing the information intervention

In this section, we discuss issues related to the design of the information intervention. First, we highlight different types of information that have been provided in prior work. Second, we discuss which sources of information are commonly used. Third, the section reviews issues related to the presentation of the information. Fourth, we lay out ways in which researchers can more credibly identify the effects of information rather than the effects of priming individuals on an issue. Finally, we review commonly used methods that employ probabilistic information treatments.

4.1 Types of information

Quantitative information Many survey experiments provide respondents with quantitative information, such as statistics based on official census data (Bottan and Perez-Truglia, 2020b; Grigorieff et al., 2020; Kuziemko et al., 2015; Roth et al., 2021b) or expert forecasts about the future of the economy (Armantier et al., 2016; Coibion et al., 2019b; Roth and Wohlfart, 2020). While quantitative information may be hard to understand for a large fraction of the population, it often facilitates the interpretation of experimental findings in the context of a theoretical framework. Moreover, together with elicited priors and posteriors numerical information allows for the calculation of learning rates (see section 8). Many times researchers provide statistical information about the behavior of others (Allcott, 2011; Coffman et al., 2017; Duflo and Saez, 2003). A commonly used strategy provides a random subset of respondents with information about others' effort choices (Cantoni et al., 2019; Hager et al., 2020, 2021) or others' beliefs, preferences and actions (Bursztyn et al., 2020b,d; Coibion et al., 2020e).

Anecdotal evidence, stories, and narratives Another highly relevant and important, but different type of information relies on qualitative anecdotes, stories or narratives.[9] This information is not based on statistics, but instead provides qualitative information which closely resembles case studies. Experiments systematically studying the role of stories, anecdotal evidence and narratives are still very scarce, and we believe a fruitful area for future research. Anecdotal information can also be communicated via pictures and videos, which may be more effective in conveying information. A literature in development economics has studied how inspirational videos change people's beliefs and economic behavior (Bernard et al., 2014; Riley, 2017).[10]

[9] Bénabou et al. (2018) study the role of narratives from a theoretical perspective.

[10] This is also related to a literature studying how the media affects people's beliefs and their behavior (Banerjee et al., 2019; Bursztyn et al., 2020a,c; DellaVigna and Kaplan, 2007; La Ferrara et al., 2012; Martinez-Bravo and Stegmann, 2017; Yanagizawa-Drott, 2014).

Tailored versus general information One key difference across information treatments is whether the treatments are tailored to individuals or whether they concern more general pieces of information. For example, while Roth and Wohlfart (2020) provide all respondents with information about general economic conditions, Roth et al. (2021b) provide individuals with economic information based on their personal characteristics. In the context of health behaviors, Prina and Royer (2014) study the impact of providing tailored information to parents about the body weight of their own school-aged children.

4.2 Sources of information

There are many different sources for information that prior research has used to exogenously vary respondents' beliefs and expectations. Researchers commonly provide respondents with official government statistics (for instance, about the unemployment rate among immigrants (Grigorieff et al., 2020)) and research evidence (for instance, about the labor market effects of immigration (Haaland and Roth, 2020), racial discrimination (Haaland and Roth, 2021), intergenerational mobility (Alesina et al., 2018c), or economic cost of pandemic response measures (Settele and Shupe, 2020)). In the context of forward-looking expectations, one method to exogenously vary expectations is the provision of expert forecasts. In the context of macroeconomic forecasts, Roth and Wohlfart (2020) provide respondents with different forecasts about the likelihood of a recession and Hager et al. (2019) provide different expert forecasts about the anticipated turnout to different protests. In experiments that aim to change perceptions of social norms, researchers provide respondents with information about the views of respondents as measured in other surveys (Bursztyn et al., 2020b). Moreover, researchers have also explored the effects of randomly providing news articles or statements from policymakers on people's beliefs and expectations (Coibion et al., 2019b). In general, it is important to consider how credible respondents find the source of information. Rafkin et al. (2021) randomize exposure to information that highlights the government's inconsistency in the context of the coronavirus epidemic. They show that when inconsistency is salient, participants have a reduced propensity to revise prior beliefs about death counts and lower self-reported trust in the government.

Generating a first stage on beliefs Sometimes, the researcher can choose between several different truthful sources of information that might differ in how closely aligned they are likely to be with people's prior beliefs. If the goal of the information provision experiment is to generate the largest possible first-stage effect on beliefs, one needs to provide information that is sufficiently different from people's prior beliefs to generate an effect. However, if the information provided is too extreme, respondents might find the information less credible (Gentzkow and Shapiro, 2006), making it necessary to strike a trade-off between providing a large information shock while retaining trust in the information provided.

4.3 Presentation of the information

How should researchers present the information in order to maximize the effectiveness of the information intervention? To minimize concerns about demand effects, the treatment should ideally be short and neutrally framed. At the same time, to generate a strong first-stage effect on beliefs, it is important to present the information in a way that maximizes understanding among respondents. One way to increase the understanding of the treatment message is to supplement the text with a graphical illustration of the information. In designs in which researchers elicit prior beliefs, an intuitive way of presenting the information graphically contrasts prior beliefs with the value from the information treatment (see, for instance, Roth and Wohlfart, 2020).

4.4 Identity of the sender

A key question in the design of the information intervention concerns the identity of the sender of the information. The identity of the sender plays a particularly important role in domains in which trust in the information is essential. For example, Alsan and Eichmeyer (2021) study persuasion regarding the medical benefits of influenza vaccination by experimentally varying race concordance between sender and receiver. They find that race concordance improves ratings of the sender and signal, but only among Black respondents. Banerjee et al. (2020) study messaging about the coronavirus pandemic using a prominent Noble laureate as the messenger. Alatas et al. (2020) study why messaging from celebrities affects the effectiveness of information dissemination on social media. Korlyakova (2021) varies whether people receive information about ethnic discrimination from experts or from ordinary people and finds larger belief updating from information provided by experts. D'Acunto et al. (2021) show that diverse policy committees are more effective in managing expectations of underrepresented groups.

In the political domain, it is possible that recipients of information will think that the information source is biased (for example, households who expect the government to manipulate official inflation statistics). Cavallo et al. (2016) show that households update their beliefs taking the perceived bias of the information source into account. In the context of macroeconomic expectation formation, Coibion et al. (2019b) study how varying the source of information about monetary policy affects updating of inflation expectations. In general, it is good practice to include direct questions on how credible and accurate people found the provided information at the end of the survey.

4.5 Priming versus information

One key challenge in information experiments is to disentangle the effects of priming from genuine belief updating.[11] Common methods to mitigate concerns about priming include (i) eliciting prior beliefs of respondents in both the treatment and the control group, (ii) separate the information provision from the main outcomes with follow-up studies, and (iii) to include an active control group (that is, the control group also receives (differential) information). The first approach guarantees that both respondents in the treatment and the control group are primed on the issue of interest. Moreover, eliciting priors allows researchers to examine whether treatment effects are stronger among respondents whose priors are less aligned with the information, which is often interpreted as evidence of genuine changes in beliefs (Armantier et al., 2016; Lergetporer et al., 2018a; Roth et al., 2021a). The second approach ensures that any short-lived priming effects are no longer relevant when the main outcomes are elicited. The third approach ensures that respondents across all conditions receive information on the issue of interest, but the information differs in terms of its content. In the following, we discuss the use of active control groups in

[11] For an excellent review on priming in economics, see Cohn and Maréchal (2016).

more depth.

Active versus passive control Many information provision experiments measure prior beliefs on an issue and then provide the treatment group with information on that issue, while a pure control group receives no information at all. An alternative design would measure prior beliefs and then provide the treatment and control group with different information (this approach of using an active control group was pioneered by Bottan and Perez-Truglia (2020b); for other recent examples of papers implementing active control groups in information provision experiments, see Akesson et al. (2020); Hager et al. (2019); Link et al. (2021); Roth and Wohlfart (2020); Roth et al. (2021b); Settele (2020)).

Providing the control group with information has several advantages for studying the causal effect of expectations on behavior. In a design with a pure control group, the variation hinges on prior beliefs. The identification mostly comes from individuals with larger misperceptions ex-ante. An active control group design generates variation in the relevant belief also among individuals with more accurate priors and therefore identifies average causal effects of beliefs on outcomes for a broader population. Moreover, receiving an information treatment may have side effects, such as uncertainty reduction, attention, and emotional responses (especially in designs where respondents have been corrected). Such side effects should arguably be similar across groups that receive different pieces of information. Finally, since prior beliefs may be measured with error and correlated with both observables and unobservables, causal identification and the interpretation of heterogeneous treatment effects are more difficult in designs with a pure control group.

There are also some advantages to having a pure control group. First, having a pure control group makes it easier to interpret correlations between the pre-treatment beliefs and the outcome of interest as beliefs among control group respondents are not affected by the treatment. Second, sometimes the policy relevant question of interest is concerned with the effect of providing a particular piece of information compared to not providing this information. See a discussion of these issues in Roth and Wohlfart (2020) in the context of experiments on macroeconomic expectations or in Hager et al. (2019) in the context of strategic interactions in political movements. Furthermore, sometimes it is not possible to have an active control group without deceiving respondents, in which case it is better to have a pure control group or employ a probabilistic design as discussed below.

4.6 Probabilistic information treatments

Researchers have started to use probabilistic information treatments to compare belief updating to Bayesian benchmarks (Eil and Rao, 2011; Mobius et al., 2015; Thaler, 2021; Zimmermann, 2020). In probabilistic information treatments, respondents are told that with a probability p they will learn the truth about a fact, and with probability $(1 - p)$ they will learn the opposite of the truth. Employing probabilistic information treatments provides researchers with fully exogenous variation in beliefs in settings where only one piece of truthful information about a benchmark is available and otherwise one would have to revert to a design with one treatment group and a control group. It also provides researchers with a Bayesian benchmark for the belief updating. However, probabilistic signals introduce more scope for motivated beliefs into the updating process, which could in turn lower the effectiveness of the information treatment (Eil and Rao, 2011; Mobius et al., 2015; Thaler, 2021). Probabilistic information treatments are usually applied to study motivated reasoning in belief updating, rather than studying causal effects of information and beliefs on behaviors. A downside of probabilistic information treatments is that they are more artificial and less natural for respondents.

Finally, Schwardmann and Van der Weele (2019) use probabilistic information treatments to study how self-confidence affects the persuasiveness of the messenger and find that higher self-confidence causally increases the persuasiveness of the messenger.

5 Measuring belief updating

In order to understand the mechanisms through which an information treatment operates it is essential to measure a rich set of beliefs that capture the theoretical mechanisms that may be at play. We first discuss how to circumvent issues related to numerical anchoring. Second, we argue that measurement of beliefs about the provided information should be more commonly used to better understand and interpret the effects of information.

Numerical anchoring An additional methodological concern for quantitative outcome measures elicited after the information provision, such as posterior beliefs about the statistic, is unconscious numerical anchoring. There are several best-practices for alleviating concerns about numerical anchoring. First, one can provide irrelevant numerical anchors and test their effects on the posterior belief of interest in order to gauge the importance of such anchoring (Cavallo et al., 2017; Coibion et al., 2020f; Roth and Wohlfart, 2020). Second, one should measure at least some quantitative beliefs on a scale that differs from the scale on which the information is communicated. Third, one should also employ qualitative measures of beliefs, which are naturally immune to numerical anchoring.

Follow-up surveys Follow-up surveys, conducted a few weeks after the initial information intervention, are an important tool used to mitigate concerns about numerical anchoring, which is a short-lived phenomenon. Follow-up surveys also alleviate concerns about consistency bias in survey responses (Falk and Zimmermann, 2012). Follow-up surveys to study whether information provision has persistent effects on beliefs, preferences, and behaviors are increasingly common and were pioneered by Kuziemko et al. (2015), Cavallo et al. (2017) and Coppock (2016) in the context of survey experiments. Follow-ups in the context of information experiments usually take place one to eight weeks after the initial information provision. An exception is Fehr et al. (2019) whose follow-up takes place one year after the initial information provision. In choosing the time between main and follow-up surveys, researchers often face a trade-off between testing for persistence and maximizing the recontact rate of respondents.

Measuring beliefs about the information Finally, in order to obtain a better understanding of the effects of the information treatment, we think that researchers should measure trust in and other beliefs about the provided information. For example, Haaland and Roth (2020) elicit a rich set of beliefs about the research evidence provided to respondents. Naturally, such explicit questions may induce significant experimenter demand effects. One way to mitigate concerns about such experimenter demand effects is to elicit incentivized measures of willingness to pay for the information of interest (Fehr et al., 2019; Haaland

and Roth, 2021; Hjort et al., 2019; Hoffman, 2016).

Cross-learning Another recurring issue in information provision experiments is cross-learning. Specifically, respondents may not only update beliefs about the object of interest but at the same time change their beliefs about other variables. For instance, Coibion et al. (2019a) find that the provision of information about inflation not only changes respondents' inflation expectations but also their beliefs about GDP growth. On the one hand, such cross-learning can be seen as a natural by-product of experimental changes in beliefs, as changes in beliefs due to natural variation are similarly often correlated across variables. On the other hand, cross-learning can complicate the interpretation of instrumental variables (IV) estimates exploiting randomized information provision, as such estimates are often compared to theoretical benchmarks which do not account for cross-learning. In other words, in the presence of substantial cross-learning it is less straightforward to interpret the effects of information on behavior through the lens of belief changes.

One way to overcome the issue of cross-learning is to hold fixed beliefs about other variables by providing identical information about other variables to respondents in both the control and the treatment groups. However, simultaneous provision of several pieces of information might arguably reduce attention to the main piece of information and lead to a weaker first stage. In any case, researchers should include measures for beliefs about other variables which could be shifted by the treatment in their survey in order to be able to detect cross-learning and to gauge its extent and implications. For instance, Link et al. (2021) provide information about future nominal interest rates and include measures of posterior inflation expectations to quantify updating of expectations about real interest rates.

6 Dealing with experimenter demand effects

One concern with information provision experiments are demand effects (de Quidt et al., 2018; Mummolo and Peterson, 2019; Zizzo, 2010).[12] While recent empirical evidence suggests a limited quantitative importance of experimenter demand effects in online surveys in some domains (de Quidt et al., 2018; Mummolo and Peterson, 2019), it is still possible that in other contexts treatment effects are confounded by experimenter demand effects as people in the different treatment arms may make differential inference about the experimenter's expectations.[13] In this section, we outline best-practice recommendations to mitigate concerns about experimenter demand effects.

Obfuscated follow-ups Haaland and Roth (2020, 2021) propose the use of obfuscated follow-ups to mitigate concerns about experimenter demand effects. Obfuscated follow-up surveys are follow-up studies with the same respondents as in the initial experiment, which are presented as an independent study to participants. Since no treatment is administered in the follow-up study, differential experimenter demand between the treatment and control group is unlikely to be a concern unless respondents nonetheless realize that the follow-up is connected to the main study. Haaland and Roth (2020, 2021) take several steps to hide the connection between their main study and their obfuscated follow-up study.

[12] In the case of surveys administered by enumerators, Kerwin and Reynoso (2021) show that reported beliefs are significantly related to interviewer knowledge and suggest corrections from the perspectives of interviewer recruitment, survey design, and experiment setup.

[13] It is plausible that the willingness to please the experimenter could vary across different decision-making domains which might increase the relevance of demand effects.

First, they collaborated with a market research company where respondents regularly receive invitations to participate in surveys. The marketing company sent generic invitations that only reveal information about pay and expected completion time. Second, they employed two different consent forms for the two surveys. Third, to give the impression that the follow-up is an independent study, they first ask respondents a series of questions about their demographics. Fourth, to further obfuscate the purpose of the follow-up, they pose questions about unrelated issues before asking any of the actual questions of interest. Following the approach proposed by Haaland and Roth (2020, 2021), Settele (2020) uses an obfuscated follow-up survey in the context of attitudes towards affirmative action.

Anonymity Anonymity has been argued to be a powerful tool against experimenter demand effects in experimental research (Hoffman et al., 1994). In the context of policy preference experiments, researchers have recently relied on the use of anonymous online petitions in order to mitigate concerns about experimenter demand effects (Grigorieff et al., 2020). A commonly used additional tool are "list methods" which aim to veil the answers of individual respondents and are increasingly applied throughout the social sciences (Bursztyn et al., 2020b; Chen and Yang, 2019; Coffman et al., 2016; Lergetporer et al., 2017).

Incentivized outcomes Over the last few years, researchers have started using incentivized outcomes in the context of survey experiments. A commonly used approach is to elicit incentivized donations to political organizations which capture specific policy preferences (Bursztyn et al., 2020d; Grigorieff et al., 2020; Haaland and Roth, 2021; Roth et al., 2021a; Settele, 2020). Presumably, demand effects should be lower in tasks in which real money is at stake.

Field outcomes A small number of studies manage to link information provision with natural outcomes from the field, such as the take-up of job offers (Bursztyn et al., 2020b), the repayment of credit card debt (Bursztyn et al., 2019), welfare take-up (Finkelstein and Notowidigdo, 2019), policy choices of politicians (Hjort et al., 2019), campaign donations (Perez-Truglia and Cruces, 2017), voting behavior (Cruz et al., 2018; Gerber et al., 2020; Kendall et al., 2015), canvassing activity using an online application (Hager et al., 2020, 2021), home sales (Bottan and Perez-Truglia, 2020a), credit card spending (Galashin et al., 2020), or stock trading choices of retail investors (Laudenbach et al., 2021). The key advantage of these studies is that they provide unobtrusive behavioral outcome data from a natural setting. Experimenter demand effects are of no concern in many of these natural settings as respondents are often not aware of the fact that they are part of an experiment. In general, given that decisions in the field involve much higher stakes than survey responses, it is unlikely that changes in these outcomes reflect demand effects.

Neutral framing How should researchers frame the information treatments? One way to minimize the relevance of experimenter demand effects is to adopt a neutral framing of the experimental instructions. The neutral framing of instructions usually makes the purpose of the experiment less transparent and draws less attention of respondents to the expectations and wishes of the experimenter. For example, Bursztyn et al. (2020d) truthfully tell respondents that they will be assigned to decide whether to authorize a donation to either a pro-immigrant or an anti-immigrant organization. This reduces concerns that researchers are perceived as politically biased.

Obfuscated information treatments One way to mitigate experimenter demand effects is to obfuscate the information treatments. Specifically, researchers can try to obfuscate the purpose of the study by providing respondents with additional pieces of information which are irrelevant, or by giving respondents tasks that give the impression that the purpose of the study is completely unrelated to the actual goal. One possibility is to give people an unrelated reason for why they receive the information of interest. For instance, researchers could tell respondents that they need to proofread or summarize pieces of information. For an example in the context of immigration attitudes, see Facchini et al. (2016). Furthermore, in experiments in which the researcher elicits incentivized prior beliefs, the purpose of the information treatment may be naturally concealed by framing the information treatment as feedback on whether the respondent's answer qualified for an extra payment.

Demand treatments de Quidt et al. (2018) propose the use of demand treatments in order to measure the sensitivity of behavior and self-reports with respect to explicit signals about the experimenter's expectations. For example, they tell respondents that they "expect that participants who are shown these instructions" will act in a particular way. The idea behind their approach is that one can use explicit signals of experimenters' wishes in order to bound the natural action. Roth and Wohlfart (2020) and Mummolo and Peterson (2019) apply demand treatments in the context of survey experiments on macroeconomic expectations and in political science, respectively, and confirm the finding that responsiveness to demand treatments is quite moderate.

Measuring beliefs about the study purpose Many research studies in economics and psychology measure beliefs about the study purpose. Demand effects are less likely a concern in an experiment or survey if participants cannot identify the intent of the study. Allcott and Taubinsky (2015) measure perceptions of study intent and show that that there is strong dispersion in perceived intent within treatment groups, suggesting that it is unclear in which way demand effects might affect behavior. To test whether respondents across treatment arms hold different beliefs about the study purpose, Bursztyn et al. (2020e) use a machine learning classifier to predict treatment status based on open-ended text responses about perceived study purpose.

Heterogeneity by self-monitoring scale Allcott and Taubinsky (2015) argue that if demand effects are driving behavior in experiments, then they should be more pronounced for respondents who are more able to detect the intent of the study and are more willing to change their choices given the experimenter's intent. Allcott and Taubinsky (2015) employ the self-monitoring scale by Snyder (1974) and find no evidence that self-monitoring ability moderates the treatment effect.

Summary Overall, evidence suggests that demand effects may be of less limited quantitative importance in online experiments in some domains (de Quidt et al., 2018). However, the importance of demand effects could vary a lot across settings. We believe that they might be a concern particularly in sensitive domains in which participants care about pleasing the experimenter, while they are less of a concern in domains in which participants care about expressing their true preferences. Since it is ex-ante unclear how relevant demand effects are, it is best-practice to include some of the above outlined checks.

7 Samples

In this section, we first provide an overview of samples that are commonly used to conduct information provision experiments, with a particular focus on the United States. We then then provide recommendations on how to measure attention in online surveys to ensure high-quality responses.

7.1 Online panels

We now discuss the advantages and disadvantages of three different types of online samples that are commonly used for conducting information provision experiments: (i) probability-based samples, (ii) online panels representative in terms of observables, and (iii) online labor markets, such as Amazon Mechanical Turk.

Probability-based samples The most representative samples are probability-based panels. In a probability-based panel, the survey company recruits the sample by randomly selecting households from a representative sample frame. People cannot join the panel unless they have been randomly selected for participation. Random sampling from a representative sample gives these panels some desirable theoretical properties relating to unbiasedness and quantifiable margins of error. However, given that the non-response rate for probability-based panels is often quite high, there is still a strong element of self-selection into the panels.[14] A clear advantage of probability-based samples is that they include more respondents who are typically under-represented in non-probability-based samples, such as low-income and rural respondents as well as respondents from the non-Internet population. The disadvantages of probability-based samples are that they are typically much costlier than convenience samples and that they typically offer the least degree of flexibility in survey design and implementation.

In the United States, a widely used probability-based panel is AmeriSpeak by NORC at the University of Chicago. The panel uses NORC's National Frame, which is designed to provide at least 97 percent sample coverage of the US population. The NORC National Frame is used for several landmark studies in the US, including the General Social Survey, which is one of the most frequently analyzed data sets in the social sciences. Other probability-based samples of the US population open to academic researchers include the RAND American Life Panel, the Understanding America Study at the University of Southern California, and the Ipsos KnowledgePanel (formerly administered by GfK).

Representative online panels Representative online panels are constructed to be representative of the general population in terms of observable characteristics, but do not use random sampling to recruit respondents. In a representative online panel, the survey company recruits respondents through, for instance, advertisements, and anyone who wants to join the panel is free to do so. The main advantage of these panels is that they are much more affordable than probability-based panels while retaining representatives

[14] In the United States, a typical response rate for probability-based samples is between 5 and 15 percent; see `https://www.pewresearch.org/methods/u-s-survey-research/our-survey-methodology-in-detail/` (accessed April 9, 2021)

in terms of some important observable characteristics, such as age, income, race, and gender. The main drawbacks of these panels are that the lack of random sampling makes it difficult to estimate the margin of error for the general population and that they do not include respondents from the non-Internet population. However, given that most probability-based panels have relatively high non-response rates, the differences in the extent of selection between probability-based samples and representative online panels might not be that large in practice.

Three large providers that are widely used in the social sciences are Dynata (formerly Research Now and Survey Sampling International), Lucid, and YouGov. While some providers (such as YouGov) aim to match higher-dimensional cells of the population (such as age X gender), others (such as Lucid) approximate marginal distributions of basic demographics in the population. Furthermore, they allow for the use of obfuscated follow-up studies. The main disadvantage of these panels is that inferences may be less externally valid and there is a concern that respondents who self-select into online panels are very different from the broader population. However, using German data, Grewenig et al. (2018) show that the online and the offline population hardly differ in terms of survey responses in the context of political views and opinions, once the survey method and observable respondent characteristics are controlled for.

Coppock and McClellan (2019) find that samples from Lucid score similarly to respondents in the American National Election Study (ANES) on the Big-5 personality inventory, show similar levels of political knowledge, and recover framing effects similar to the ones observed in a probability-based sample (the General Social Survey). Haaland and Roth (2021) find similar experimental results using a sample from a representative online panel provider and a probability-based sample. Other comparable providers are Respondi, Prolific, and the Qualtrics panel.

Amazon Mechanical Turk The third type of available online sample are online labor markets, such as Amazon Mechanical Turk (MTurk), which are widely used in the social sciences and economics (Kuziemko et al., 2015). Coppock (2018) conducts 15 replication experiments and finds a very high degree of replicability of survey experiments in the field of political science with MTurk as compared to nationally representative samples. Horton et al. (2011) replicate several well-known lab experiments using MTurk, concluding that online experiments on MTurk are just as valid as traditional physical lab experiments. However, recent studies suggest that data quality on MTurk has been declining over time, partly through the proliferation of bots (automated computer programs) and non-serious respondents, which threatens the data quality on the platform if sensible screening procedures are not implemented (Ahler et al., 2019; Chmielewski and Kucker, 2020). To maximize data quality on MTurk, one should only allow workers that have completed a large number of previous tasks with a high completion rate. Furthermore, in the actual survey, one should include fraud detection tools to rule out bots, such as a CAPTCHA, at the beginning of the survey. While MTurk is less representative than most other survey platforms, the platform has some important advantages. First, data collection speed is typically very fast and it offers researchers maximum flexibility in terms of research design. Second, since users sign up for MTurk with their own credit card, it is also possible to incentivize respondents with real money (respondents from more representative panel platforms are typically paid in panel currencies that can be converted into gift vouchers).

Third, it is possible to conduct follow-up studies with low attrition rates (Grigorieff et al., 2020).

7.2 Measuring attention in online surveys

Screeners One concern in online surveys is that respondents are inattentive and speed through the surveys (Krosnick, 1991). We recommend using multiple attention checks in online surveys. Recent research suggests that the inclusion of attention checks does not affect estimated treatment effects, but it allows researchers to study how measured attention affects behavior (Berinsky et al., 2014; Kane and Barabas, 2019). One example of an attention screener is the following:

> The next question is about the following problem. In questionnaires like ours, sometimes there are participants who do not carefully read the questions and just quickly click through the survey. This means that there are a lot of random answers which compromise the results of research studies. To show that you read our questions carefully, please enter turquoise as your answer to the next question. What is your favorite color?

There are at least two features of attention checks that we consider important: first, it is important for attention checks to explain to participants why researchers use these attention checks. This explanation can mitigate concerns about negative emotional reactions to the use of attention checks on the part of participants. Second, we think that attention checks should be simple to understand and should not be too cognitively demanding. Therefore, having an unambiguous and easy-to-understand question is important. For an excellent review on attention checks, see Berinsky et al. (2014).

Open-ended questions Bots have been identified as a threat to online surveys. On top of standard bot protections, such as asking respondents to categorize distorted pictures that computers cannot easily recognize (CAPTCHAs), we recommend using at least two open-ended questions in the survey, e.g. to inquire about feedback on the survey or to ask about the study purpose. These open-ended questions are a useful tool to assess data quality and to identify bots that may provide identical (and/or non-sensical) responses to different open-ended questions.

8 Typical effect sizes and recommended sample sizes

In this section, we briefly discuss typical effect sizes from information provision experiments.

Learning rates Information experiments usually measure belief updating using either qualitative or quantitative questions. In the context of quantitative beliefs, papers often include a calculation of learning rates. To calculate such learning rates, we require prior beliefs about the provided piece of information.[15] Moreover, typically we observe both a treatment group which receives information and a control group, which does not receive any information. To quantify the extent to which the respondents update their beliefs towards the signal they receive during the information treatment, one can estimate the

[15] An exception are designs with an active control group, in which the average learning rate can be inferred from comparing the difference in posteriors between treatment groups with the difference in the provided signals.

following specification:

$$\text{Updating}_i = \beta_0 + \beta_1 \text{Treatment}_i \text{Perc.-gap}_i + \beta_2 \text{Treatment}_i + \beta_3 \text{Perc.-gap}_i + \varepsilon_i$$

where Updating_i is defined as the difference between the respondent's posterior and prior about the quantity of interest. Since priors about the quantity of interest should be balanced across treatment arms, one could alternatively directly use the posterior as left-hand-side variable. The perception gap, Perc.-gap_i, is the difference between the true signal and the respondent's prior belief about the signal. The key coefficient of interest, β_1, captures the extent of belief updating toward the provided signal among respondents in the treatment group, on top of any updating that also happens for respondents in the control group. β_2 captures the average treatment effect on respondents' beliefs to the extent it does not depend on individual priors. β_3 measures the extent to which changes in beliefs in the control group depend on the perception gap. It is essential to control for Perc.-gap_i in a non-interacted form, as also respondents who were not provided with the information may change beliefs into the direction of the signal, e.g. because they have thought more carefully about the question once they are asked a second time, or because they have committed a typo the first time they stated their beliefs (Fuster et al., 2020).

Cavallo et al. (2017) and Cullen and Perez-Truglia (2018) show how the coefficient β_1 can be given a more structural interpretation. Specifically, under Bayesian updating with normally distributed priors and signals (where the moments of these distributions are independent of each other), and quadratic loss function, updating of beliefs will be linear in the perception gap. β_1 captures the weight respondents put on the signal, while putting a weight of $1 - \beta_1$ on their prior belief.

Table 1 gives an overview of estimated learning rates from a few select information experiments that provide quantitative information and calculate such learning rates. Many of the papers estimating quantitative learning rates focus on macroeconomic expectations. For instance, Armantier et al. (2016) find a learning rate of 0.39 for 1-year inflation expectations in response to a professional forecast. Armona et al. (2019) estimate an instantaneous learning rate of 0.18 for house price growth in response to information about past house price growth. In a two-month follow-up, they estimate a learning rate of 0.13, indicating a high degree of persistence. Cavallo et al. (2016) estimate learning rates between 0.3 and 0.8 for inflation expectations in response to information about official inflation statistics or product price changes, which persist at about half of their initial values in a two-month follow-up. Roth and Wohlfart (2020) estimate a learning rate of 0.32 for recession expectations in response to a professional forecast. In a two-week follow-up, they document a learning rate of 0.13, indicating a moderate degree of persistence. Taken together, these papers document that people persistently learn from the information provided, but that effects in most cases become weaker over time.

Effect sizes on beliefs versus behavior Effect sizes on self-reported attitudes and behavioral measures are typically much smaller in magnitude than effect sizes on belief updating in response to information treatments. For instance, Alesina et al. (2018c) employ an information treatment to generate exogenous variation in perceptions of social mobility. While perceptions about the probability of remaining in the bottom quintile of the income distribution increase by 9.7 percentage points—thus making treated respondents substantially more pessimistic about the social mobility process—the authors find essentially no average impact on policy preferences. Similarly, an experiment by Kuziemko et al. (2015) provides respondents with accurate information about the income distribution. They find a large effect on beliefs about income inequality: treated respondents are 12 percentage points more likely to believe that income inequality has increased. By contrast, policy preferences are largely unaffected by the treatment. Haaland and Roth (2020) report results from an experiment where effect sizes on beliefs and preferences are quite similar in magnitude. Specifically, they provide respondents with research evidence showing no adverse labor market impacts of low-skilled immigration. Treated respondents become 17.1 percent of a standard deviation more optimistic about the labor market impacts of low-skilled immigrants and 14.1 percent of a standard deviation more in favor of low-skilled immigration.[16]

Instrumental variable estimation and behavioral elasticities One way to illustrate effect sizes is to estimate two-stage least squares specifications, where the endogenous belief of interest is instrumented with the randomized information provision. For example, Bottan and Perez-Truglia (2020a) find that a 1 percentage point increase in home price expectations reduces the probability of selling within 6 months by 2.5 percentage points. Roth and Wohlfart (2020) find that a 10 percentage point increase in the perceived likelihood of a recession leads to a decrease in planned consumption growth by 13 percent of a standard deviation. A simple measure of the effect of beliefs on behavior that is

[16]In some cases information interventions not only fail to correct, but even increase misperceptions among the targeted ideological group (Nyhan and Reifler, 2010). However, while the evidence on the effectiveness of correction of misperceptions in the political domain is mixed, such "backfiring" effects seem to be the exception (Guess et al., 2020; Nyhan, 2020, 2021).

comparable across settings are "behavioral elasticities." Such elasticities can be calculated by regressing the log of the outcome of interest on the log of posterior beliefs instrumented by the treatment assignment. For example, Cullen and Perez-Truglia (2018) find that increasing the perceived manager salary by 10% would increase the number of hours worked by 1.5%. The key advantage of these approaches is that they make it easier to compare results across settings. The key disadvantage is that the exclusion restriction needed for an IV estimation may not hold as the information provided may change several beliefs simultaneously (see our discussion on cross-learning).

Sample sizes While information provision experiments often produce relatively large effect sizes on beliefs, effect sizes on stated preferences or behavioral outcomes are typically much lower. Indeed, it is not uncommon to observe null effects on the main outcomes of interest despite a large and significant first stage on beliefs (e.g., Alesina et al. 2018c; Haaland and Roth 2021; Kuziemko et al. 2015). Furthermore, how elastic different outcomes are with respect to changes in beliefs naturally varies a lot across different settings, making it difficult to make a generic recommendation about optimal sample sizes. For instance, an information provision experiment studying actual voting turnout—a sticky outcome where it is unrealistic to expect large effects of an information treatment—requires a larger sample size than a similar information provision experiment studying self-reported voting intentions.

While the optimal sample size depends on the context, null findings are not uncommon in information provision experiments and it is important to have sufficient statistical power to be able to measure a null finding relatively precisely. As a minimum, we think information provision experiments should have at least 80 percent power to detect a treatment effect of 15 percent of a standard deviation. This requires a sample size of at least 700 respondents per treatment arm of interest. For studies including a follow-up study, one should take into account the likely attrition between the main study and the follow-up and adjust the power calculation accordingly. For instance, if one expects a 30 percent attrition between the main study and a follow-up, one needs an initial sample size of 1,000 respondents per treatment arm to have 80 percent power to detect a treatment effect of 15 percent of a standard deviation in the follow-up.

9 Concluding remarks

Information provision experiments are a powerful method to test economic theories and answer policy-relevant questions. As shown in Figure 1, the use of information provision experiments has grown considerably in economics over the last decade. Furthermore, as our survey of the literature illustrates, they have become popular in most sub-fields of economics. Given the importance of generating exogenous variation in beliefs to test many influential economic theories, and given the potential of information provision experiments to address questions of high policy relevance, we believe that such experiments will continue to grow in popularity. We hope that the methodological considerations and best-practice recommendations discussed in this review will contribute to this growth by lowering the barriers to conduct information provision experiments for researchers previously unfamiliar with the methodology.

Common applications of information provision experiments include studying belief formation and how exogenous changes in beliefs affect economic behavior. The literature has demonstrated that the elasticity of beliefs and preferences with respect to information varies a lot between different domains and settings. Going forward, it will be important to provide systematic evidence on what determines the effectiveness of information in changing beliefs. Similarly, it will be important to better understand why exogenous changes in beliefs lead to large changes in behavior in some domains but not in others.

For instance, factors that are likely to be important for belief updating include the strength of prior beliefs, the complexity of the information, and people's experience in processing information. There could also be a key role for attention and memory in shaping the associations that come to mind when being presented with information, which may affect learning from the information.[17] The relative importance of these different factors is currently not well understood, and they are likely of key importance for understanding differential effects of information across contexts. Furthermore, to systematically assess the relative importance of different beliefs in shaping economic behavior, we believe that information experiments that are designed to allow for a structural interpretation of estimated elasticities between beliefs and behaviors will be especially valuable going forward.

References

Abebe, Girum Tefera, Stefano Caria, Marcel Fafchamps, Paolo Falco, Simon Franklin, and Simon Quinn, "Anonymity or Distance? Job Search and Labour Market Exclusion in a Growing African City," *Review of Economic Studies*, 2020.

[17]For recent formal models of attention, see work by Bordalo et al. (2016, 2020); Gennaioli and Shleifer (2010).

Acemoglu, Daron, Ali Cheema, Asim I Khwaja, and James A Robinson, "Trust in State and Non-State Actors: Evidence from Dispute Resolution in Pakistan," *Journal of Political Economy*, 2020, *128* (8), 3090–3147.

Ahler, Douglas J and Gaurav Sood, "The Parties in our Heads: Misperceptions about Party Composition and their Consequences," *Journal of Politics*, 2018, *80* (3), 964–981.

_ , Carolyn E Roush, and Gaurav Sood, "The Micro-Task Market for Lemons: Data Quality on Amazon's Mechanical Turk," 2019.

Ajayi, Kehinde F, Willa H Friedman, and Adrienne M Lucas, "The importance of information targeting for school choice," *American Economic Review*, 2017, *107* (5), 638–43.

Aker, Jenny C, Paul Collier, and Pedro C Vicente, "Is information power? Using mobile phones and free newspapers during an election in Mozambique," *Review of Economics and Statistics*, 2017, *99* (2), 185–200.

Akesson, Jesper, Sam Ashworth-Hayes, Robert Hahn, Robert D Metcalfe, and Itzhak Rasooly, "Fatalism, Beliefs, and Behaviors During the COVID-19 Pandemic," Working Paper 27245, National Bureau of Economic Research May 2020.

Alatas, Vivi, Arun G Chandrasekhar, Markus Mobius, Benjamin A Olken, and Cindy Paladines, "Designing Effective Celebrity Public Health Messaging: Results From A Nationawide Twitter Experiment in Indonesia," 2020.

Alesina, Alberto, Armando Miano, and Stefanie Stantcheva, "Immigration and redistribution," Technical Report, National Bureau of Economic Research 2018.

_ , Michela Carlana, Eliana La Ferrara, and Paolo Pinotti, "Revealing Stereotypes: Evidence from immigrants in schools," Technical Report, National Bureau of Economic Research 2018.

_ , Stefanie Stantcheva, and Edoardo Teso, "Intergenerational Mobility and Preferences for Redistribution," *American Economic Review*, 2018, *108* (2), 521–554.

Allcott, Hunt, "Social norms and energy conservation," *Journal of Public Economics*, 2011, *95* (9-10), 1082–1095.

_ , "The welfare effects of misperceived product costs: Data and calibrations from the automobile market," *American Economic Journal: Economic Policy*, 2013, *5* (3), 30–66.

_ **and Christopher Knittel**, "Are consumers poorly informed about fuel economy? Evidence from two experiments," *American Economic Journal: Economic Policy*, 2019, *11* (1), 1–37.

_ **and Dmitry Taubinsky**, "Evaluating behaviorally motivated policy: Experimental evidence from the lightbulb market," *American Economic Review*, 2015, *105* (8), 2501–38.

_ **, Levi Boxell, Jacob Conway, Matthew Gentzkow, Michael Thaler, and David Yang**, "Polarization and public health: Partisan differences in social distancing during the coronavirus pandemic," *Journal of Public Economics*, 2020, *191*, 104254.

Alsan, Marcella and Sarah Eichmeyer, "Persuasion in Medicine: Messaging to Increase Vaccine Demand," Technical Report, National Bureau of Economic Research 2021.

Altmann, Steffen, Armin Falk, Simon Jäger, and Florian Zimmermann, "Learning about job search: A field experiment with job seekers in Germany," *Journal of Public Economics*, 2018, *164*, 33–49.

Andrabi, Tahir, Jishnu Das, and Asim Ijaz Khwaja, "Report cards: The impact of providing school and child test scores on educational markets," *American Economic Review*, 2017, *107* (6), 1535–1563.

Andre, Peter, Carlo Pizzinelli, Christopher Roth, and Johannes Wohlfart, "Subjective Models of the Macroeconomy: Evidence from Experts and a Representative Sample," *Available at SSRN 3355356*, 2019.

Ansolabehere, Stephen, Marc Meredith, and Erik Snowberg, "Asking about numbers: Why and how," *Political Analysis*, 2013, *21* (1), 48–69.

Armand, Alex, Alexander Coutts, Pedro C Vicente, and Inês Vilela, "Does information break the political resource curse? Experimental evidence from Mozambique," *American Economic Review*, 2020, *110* (11), 3431–53.

Armantier, Olivier, Scott Nelson, Giorgio Topa, Wilbert van der Klaauw, and Basit Zafar, "The Price Is Right: Updating Inflation Expectations in a Randomized Price Information Experiment," *Review of Economics and Statistics*, 2016, *98* (3), 503–523.

_ , Wändi Bruine de Bruin, Giorgio Topa, Wilbert Klaauw, and Basit Zafar, "Inflation Expectations and Behavior: Do Survey Respondents Act on their Beliefs?," *International Economic Review*, 2015, *56* (2), 505–536.

Armona, Luis C, Andreas Fuster, and Basit Zafar, "Home Price Expectations and Behavior: Evidence from a Randomized Information Experiment," *Reviw of Economic Studies*, 2019.

Attanasio, Orazio, Teodora Boneva, and Christopher Rauh, "Parental Beliefs about Returns to Different Types of Investments in School Children," *Journal of Human Resources*, 2021.

Banerjee, Abhijit, Eliana La Ferrara, and Victor H Orozco-Olvera, "The entertaining way to behavioral change: Fighting HIV with MTV," 2019.

_ , Marcella Alsan, Emily Breza, Arun G Chandrasekhar, Abhijit Chowdhury, Esther Duflo, Paul Goldsmith-Pinkham, and Benjamin A Olken, "Messages on covid-19 prevention in india increased symptoms reporting and adherence to preventive behaviors among 25 million recipients with similar effects on non-recipient members of their communities," Technical Report, National Bureau of Economic Research 2020.

_ , Rema Hanna, Jordan Kyle, Benjamin A Olken, and Sudarno Sumarto, "Tangible information and citizen empowerment: Identification cards and food subsidy programs in Indonesia," *Journal of Political Economy*, 2018, *126* (2), 451–491.

— , **Sharon Barnhardt, and Esther Duflo**, "Movies, margins, and marketing: Encouraging the adoption of iron-fortified salt," in "Insights in the Economics of Aging," University of Chicago Press, 2015, pp. 285–306.

Bansak, Kirk, Jens Hainmueller, and Dominik Hangartner, "How economic, humanitarian, and religious concerns shape European attitudes toward asylum seekers," *Science*, 2016, *354* (6309), 217–222.

Barari, Soubhik, Stefano Caria, Antonio Davola, Paolo Falco, Thiemo Fetzer, Stefano Fiorin, Lukas Hensel, Andriy Ivchenko, Jon Jachimowicz, Gary King et al., "Evaluating COVID-19 public health messaging in Italy: Self-reported compliance and growing mental health concerns," *medRxiv*, 2020.

Barrera, Oscar, Sergei Guriev, Emeric Henry, and Ekaterina Zhuravskaya, "Facts, alternative facts, and fact checking in times of post-truth politics," *Journal of Public Economics*, 2020, *182*, 104123.

Barron, Kai, "Belief updating: does the 'good-news, bad-news' asymmetry extend to purely financial domains?," *Experimental Economics*, 2021, *24* (1), 31–58.

Baseler, Travis, "Hidden Income and the Perceived Returns to Migration: Experimental Evidence from Kenya," *Available at SSRN 3534715*, 2021.

Belot, Michele, Philipp Kircher, and Paul Muller, "How wage announcements affect job search-a field experiment," 2018.

— , — , **and** — , "Providing advice to jobseekers at low cost: An experimental study on online advice," *Review of Economic Studies*, 2019, *86* (4), 1411–1447.

Bénabou, Roland, Armin Falk, and Jean Tirole, "Narratives, imperatives, and moral reasoning," Technical Report, National Bureau of Economic Research 2018.

Benjamin, Daniel J, Don A Moore, and Matthew Rabin, "Biased Beliefs About Random Samples: Evidence from Two Integrated Experiments," Working Paper 23927, National Bureau of Economic Research October 2017.

Bennear, Lori, Alessandro Tarozzi, Alexander Pfaff, Soumya Balasubramanya, Kazi Matin Ahmed, and Alexander Van Geen, "Impact of a randomized controlled trial in arsenic risk communication on household water-source choices in Bangladesh," *Journal of Environmental Economics and Management*, 2013, *65* (2), 225–240.

Bérgolo, Marcelo L, Rodrigo Ceni, Guillermo Cruces, Matias Giaccobasso, and Ricardo Perez-Truglia, "Tax audits as scarecrows: Evidence from a large-scale field experiment," Technical Report, National Bureau of Economic Research 2017.

Berinsky, Adam J, "Telling the Truth About Believing the Lies? Evidence for the Limited Prevalence of Expressive Survey Responding," *Journal of Politics*, 2018, *80* (1), 211–224.

_ **, Michele F Margolis, and Michael W Sances**, "Separating the Shirkers from the Workers? Making Sure Respondents Pay Attention on Self-Administered Surveys," *American Journal of Political Science*, 2014, *58* (3), 739–753.

Berkes, Jan, Frauke Peter, C Spiess, and Felix Weinhardt, "Information Provision and Postgraduate Studies," 2019.

Bernard, Tanguy, Stefan Dercon, Kate Orkin, Alemayehu Taffesse et al., "The future in mind: Aspirations and forward-looking behaviour in rural Ethiopia," 2014.

Beshears, John, James J Choi, David Laibson, Brigitte C Madrian, and Katherine L Milkman, "The effect of providing peer information on retirement savings decisions," *Journal of Finance*, 2015, *70* (3), 1161–1201.

Binder, Carola, "Coronavirus Fears and Macroeconomic Expectations," *Review of Economics and Statistics*, 10 2020, *102* (4), 721–730.

_ **and Alex Rodrigue**, "Household Informedness and Long-Run Inflation Expectations: Experimental Evidence," *Southern Economic Journal*, 2018, *85* (2), 580–598.

Bleemer, Zachary and Basit Zafar, "Intended college attendance: Evidence from an experiment on college returns and costs," *Journal of Public Economics*, 2018, *157*, 184–211.

Blesse, Sebastian, Florian Buhlmann, and Philipp Doerrenberg, "Do people really want a simple tax system? Evidence on preferences towards income tax simplification," *Working Paper*, 2019.

Bond, Timothy N and Kevin Lang, "The Sad Truth About Happiness Scales," *Journal of Political Economy*, 2019, *127* (4), 1629–1640.

Boneva, Teodora and Christopher Rauh, "Socio-economic gaps in university enrollment: The role of perceived pecuniary and non-pecuniary returns," 2017.

_ **and** _ , "Parental Beliefs about Returns to Educational Investments—The Later the Better?," *Journal of the European Economic Association*, 2018, *16* (6), 1669–1711.

_ **, Marta Golin, Christopher Rauh et al.**, "Can Perceived Returns Explain Enrollment Gaps in Postgraduate Education?," Technical Report 2019.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer, "Stereotypes," *Quarterly Journal Economics*, 2016, *131* (4), 1753–1794.

_ **, Nicola Gennaioli, and Andrei Shleifer**, "Memory, Attention, and Choice," *Quarterly Journal of Economics*, 2020, *135* (3), 1399–1442.

Bott, Kristina M, Alexander W Cappelen, Erik Ø Sørensen, and Bertil Tungodden, "You've got mail: A randomized field experiment on tax evasion," *Management Science*, 2020, *66* (7), 2801–2819.

Bottan, Nicolas L and Ricardo Perez-Truglia, "Betting on the House: Subjective Expectations and Market Choices," *Available at SSRN*, 2020.

_ **and** _ , "Choosing Your Pond: Revealed-Preference Estimates of Relative Income Concerns," *Review of Economics and Statistics*, 2020.

Brier, Glenn W, "Verification of forecasts expressed in terms of probability," *Monthly weather Review*, 1950, *78* (1), 1–3.

Bruin, Wändi Bruine De, Charles F Manski, Giorgio Topa, and Wilbert Van Der Klaauw, "Measuring Consumer Uncertainty about Future Inflation," *Journal of Applied Econometrics*, 2011, *26* (3), 454–478.

Bullock, John G, Alan S Gerber, Seth J Hill, Gregory A Huber et al., "Partisan Bias in Factual Beliefs about Politics," *Quarterly Journal of Political Science*, 2015, *10* (4), 519–578.

Bursztyn, Leonardo, Aakaash Rao, Christopher Roth, and David Yanagizawa-Drott, "Misinformation during a pandemic," *University of Chicago, Becker Friedman Institute for Economics Working Paper*, 2020, (2020-44).

_ , Alessandra L. González, and David Yanagizawa-Drott, "Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia," *American Economic Review*, October 2020, *110* (10), 2997–3029.

_ , Davide Cantoni, Patricia Funk, and Noam Yuchtman, "Identifying the Effect of Election Closeness on Voter Turnout:Evidence from Swiss Referenda," *NBER working paper*, 2020, (w23490).

_ , Florian Ederer, Bruno Ferman, and Noam Yuchtman, "Understanding Mechanisms Underlying Peer Effects: Evidence from a Field Experiment on Financial Decisions," Technical Report 4 2014.

_ , Georgy Egorov, and Stefano Fiorin, "From Extreme to Mainstream: The Erosion of Social Norms," *American Economic Review*, November 2020, *110* (11), 3522–48.

_ , Ingar Haaland, Aakaash Rao, and Christopher Roth, "Disguising Prejudice: Popular Rationales as Excuses for Intolerant Expression," NBER Working Paper 27288 May 2020.

_ , Stefano Fiorin, Daniel Gottlieb, and Martin Kanz, "Moral Incentives in Credit Card Debt Repayment: Evidence from a Field Experiment," *Journal of Political Economy*, 2019, *127* (4), 1641–1683.

Camerer, Colin F and Robin M Hogarth, "The Effects of Financial Incentives in Experiments: A Review and Capital-labor-production Framework," *Journal of risk and uncertainty*, 1999, *19* (1-3), 7–42.

Cantoni, Davide, David Y Yang, Noam Yuchtman, and Y Jane Zhang, "Protests as strategic games: experimental evidence from Hong Kong's antiauthoritarian movement," *Quarterly Journal of Economics*, 2019, *134* (2), 1021–1077.

Card, David, Alexandre Mas, Enrico Moretti, and Emmanuel Saez, "Inequality at work: The effect of peer salaries on job satisfaction," *American Economic Review*, 2012, *102* (6), 2981–3003.

Carneiro, Pedro Manuel, Lucy Kraftman, Giacomo Mason, Lucie Moore, Imran Rasul, and Molly Scott, "The Impacts of a Multifaceted Pre-natal Intervention on Human Capital Accumulation in Early Life," 2020.

Carranza, Eliana, Robert Garlick, Kate Orkin, and Neil Rankin, "Job search and hiring with two-sided limited information about workseekers' skills," Technical Report, World Bank 2020.

Cavallo, Alberto, Guillermo Cruces, and Ricardo Perez-Truglia, "Learning from potentially-biased statistics: Household inflation perceptions and expectations in Argentina," *Brooking Papers on Economic Activity*, 2016, pp. 59–108.

_ , _ , and _ , "Inflation Expectations, Learning and Supermarket Prices: Evidence from Field Experiments," *American Economic Journal: Macroeconomics*, 2017, *9* (3), 1–35.

Chen, Yan, F Maxwell Harper, Joseph Konstan, and Sherry Xin Li, "Social comparisons and contributions to online communities: A field experiment on movielens," *American Economic Review*, 2010, *100* (4), 1358–98.

_ , Fangwen Lu, and Jinan Zhang, "Social comparisons, status and driving behavior," *Journal of Public Economics*, 2017, *155*, 11–20.

Chen, Yuyu and David Y. Yang, "The Impact of Media Censorship: 1984 or Brave New World?," *American Economic Review*, 2019, *109* (6), 2294–2332.

_ , Hui Wang, and David Y Yang, "Salience of History and the Preference for Redistribution," *Available at SSRN 2717651*, 2016.

Chetty, Raj and Emmanuel Saez, "Teaching the tax code: Earnings responses to an experiment with EITC recipients," *American Economic Journal: Applied Economics*, 2013, *5* (1), 1–31.

Chmielewski, Michael and Sarah C Kucker, "An MTurk crisis? Shifts in data quality and the impact on study results," *Social Psychological and Personality Science*, 2020, *11* (4), 464–473.

Chong, Alberto, Marco Gonzalez-Navarro, Dean Karlan, and Martin Valdivia, "Effectiveness and spillovers of online sex education: Evidence from a randomized evaluation in Colombian public schools," *NBER Working Paper Series*, 2013, p. 18776.

Chopra, Felix, Ingar Haaland, and Christopher Roth, "Do People Value More Informative News?," Working Paper 2021.

Ciancio, Alberto, Adeline Delavande, Hans-Peter Kohler, and Iliana V Kohler, "Mortality Risk Information, Survival Expectations and Sexual Behaviors," 2020.

Clifford, Scott, Geoffrey Sheagley, and Spencer Piston, "Increasing Precision in Survey Experiments Without Introducing Bias," *Working Paper*, 2020.

Coffman, Katherine B, Lucas C Coffman, and Keith M Marzilli Ericson, "The Size of the LGBT Population and the Magnitude of Anti-Gay Sentiment Are Substantially Underestimated," *Management Science*, 2016, *63* (10), 3168–3186.

Coffman, Lucas C, Clayton R Featherstone, and Judd B Kessler, "Can social information affect what job you choose and keep?," *American Economic Journal: Applied Economics*, 2017, *9* (1), 96–117.

Cohn, Alain and Michel André Maréchal, "Priming in economics," *Current Opinion in Psychology*, 2016, *12*, 17–21.

Coibion, Olivier and Yuriy Gorodnichenko, "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts," *American Economic Review*, 2015, *105* (8), 2644–2678.

— , **Dimitris Georgarakos, Yuriy Gorodnichenko, and Maarten van Rooij**, "How Does Consumption Respond to News about Inflation? Field Evidence from a Randomized Control Trial," *National Bureau of Economic Research Working Paper*, 2019.

— , — , — , **and Michael Weber**, "Forward guidance and household expectations," Technical Report, National Bureau of Economic Research 2020.

— , — , — , **Geoff Kenny, and Michael Weber**, "The Effect of Macroeconomic Uncertainty on Household Spending," Technical Report, National Bureau of Economic Research 2021.

— , **Yuriy Gorodnichenko, and Michael Weber**, "Monetary Policy Communications and their Effects on Household Inflation Expectations," Technical Report, National Bureau of Economic Research 2019.

— , — , **and** — , "Does Policy Communication During COVID work?," *Chicago Booth Research Paper*, 2020, (20-15).

— , — , **and** — , "Political Polarization and Expected Economic Outcomes," Technical Report, National Bureau of Economic Research 2020.

— , — , **and** — , "Fiscal Policy and Households' Inflation Expectations: Evidence from a Randomized Control Trial," Technical Report, National Bureau of Economic Research 2021.

— , — , **and Saten Kumar**, "How Do Firms Form their Expectations? New Survey Evidence," *American Economic Review*, 2018, *108* (9), 2671–2713.

— , — , **and Tiziano Ropele**, "Inflation expectations and firm decisions: New causal evidence," *Quarterly Journal of Economics*, 2019.

— , — , **Edward S Knotek II, and Raphael Schoenle**, "Average Inflation Targeting and Household Expectations," Technical Report, National Bureau of Economic Research 2020.

— , — , **Saten Kumar, and Jane Ryngaert**, "Do You Know That I Know That You Know...? Higher-Order Beliefs in Survey Data," *Quarterly Journal of Economics*, 2020.

— , — , — , **and Mathieu Pedemonte**, "Inflation Expectations As a Policy Tool?," *Journal of International Economics*, 2020, *124*, 103297.

Conlon, John J, "Major Malfunction: A Field Experiment Correcting Undergraduates' Beliefs about Salaries," *Journal of Human Resources*, 2019, pp. 0317–8599R2.

Coppock, Alexander, "The Persistence of Survey Experimental Treatment Effects," *Working Paper*, 2016.

— , "Generalizing from Survey Experiments Conducted on Mechanical Turk: A Replication Approach," *Political Science Research and Methods*, 2018, pp. 1–16.

— **and Oliver A McClellan**, "Validating the demographic, political, psychological, and experimental results obtained from a new source of online survey respondents," *Research & Politics*, 2019, *6* (1), 1–14.

Cruces, Guillermo, Ricardo Perez-Truglia, and Martin Tetaz, "Biased Perceptions of Income Distribution and Preferences for Redistribution: Evidence from a Survey Experiment," *Journal of Public Economics*, 2013, *98*, 100–112.

Cruz, Cesi, Philip Keefer, and Julien Labonne, "Buying Informed Voters: New Effects of Information on Voters and Candidates," Technical Report 635 2021.

— , — , — , **and Francesco Trebbi**, "Making policies matter: Voter responses to campaign promises," Technical Report, National Bureau of Economic Research 2018.

Cullen, Zoë and Ricardo Perez-Truglia, "How much does your boss make? the effects of salary comparisons," Technical Report, National Bureau of Economic Research 2018.

D'Acunto, Francesco, Andreas Fuster, and Michael Weber, "Diverse Policy Committees Can Reach Underrepresented Groups," *Working Paper*, 2021.

Dafoe, Allan, Baobao Zhang, and Devin Caughey, "Information Equivalence in Survey Experiments," *Political Analysis*, 2018, *26* (4), 399–416.

Danz, David, Alistair J Wilson, and Lise Vesterlund, "Belief Elicitation: Limiting Truth Telling with Information or Incentives," 2020.

De Neve, Jan-Emmanuel, Clément Imbert, Johannes Spinnewijn, Teodora Tsankova, and Maarten Luts, "How to Improve Tax Compliance? Evidence from Population-Wide Experiments in Belgium," *Journal of Political Economy*, 2021, *129* (5), 1425–1463.

de Quidt, Jonathan, Johannes Haushofer, and Christopher Roth, "Measuring and Bounding Experimenter Demand," *American Economic Review*, 2018, *108* (11), 3266–3302.

Delavande, Adeline, "Probabilistic Expectations in Developing Countries," *Annual Review of Economics*, 2014, *6* (1), 1–20.

_ **and Basit Zafar**, "University Choice: The Role of Expected earnings, Non-pecuniary Outcomes and Financial Constraints," *Journal of Political Economy*, 2019, *127* (5).

_ **and Hans-Peter Kohler**, "Subjective Expectations in the Context of HIV/AIDS in Malawi," *Demographic Research*, 2009, *20*, 817–875.

_ **and** _ , "The impact of HIV testing on subjective expectations and risky behavior in Malawi," *Demography*, 2012, *49* (3), 1011–1036.

_ **, Xavier Giné, and David McKenzie**, "Measuring Subjective Expectations in Developing Countries: A Critical Review and New Evidence," *Journal of Development Economics*, 2011, *94* (2), 151–163.

DellaVigna, Stefano and Ethan Kaplan, "The Fox News effect: Media bias and voting," *Quarterly Journal of Economics*, 2007, *122* (3), 1187–1234.

Dizon-Ross, Rebecca, "Parents' beliefs about their children's academic ability: Implications for educational investments," *American Economic Review*, 2019, *109* (8), 2728–65.

Doerrenberg, Philipp and Andreas Peichl, "Tax morale and the role of social norms and reciprocity. Evidence from a randomized survey experiment," 2018.

— **and Jan Schmitz**, "Tax compliance and information provision – A field experiment with small firms," Discussion paper 15–028, ZEW-Centre for European Economic Research 2015.

Dolls, Mathias, Philipp Doerrenberg, Andreas Peichl, and Holger Stichnoth, "Do retirement savings increase in response to information about retirement and expected pensions?," *Journal of Public Economics*, 2018, *158*, 168–179.

Duflo, Esther and Emmanuel Saez, "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment," *Quarterly Journal of Economics*, 2003, *118* (3), 815–842.

Dupas, Pascaline, "Do teenagers respond to HIV risk information? Evidence from a field experiment in Kenya," *American Economic Journal: Applied Economics*, 2011, *3* (1), 1–34.

— **and Edward Miguel**, "Impacts and determinants of health levels in low-income countries," in "Handbook of economic field experiments," Vol. 2, Elsevier, 2017, pp. 3–93.

D'Acunto, Francesco, Daniel Hoang, Maritta Paloviita, and Michael Weber, "IQ, expectations, and choice," Technical Report, National Bureau of Economic Research 2019.

— **,** — **,** — **, and** — , "Effective Policy Communication: Targets versus Instruments," 2020.

Eil, David and Justin M Rao, "The good news-bad news effect: asymmetric processing of objective information about yourself," *American Economic Journal: Microeconomics*, 2011, *3* (2), 114–138.

Enke, Benjamin and Thomas Graeber, "Cognitive Uncertainty," Technical Report, National Bureau of Economic Research 2019.

Eriksson, Kimmo and Brent Simpson, "What do Americans know about inequality? It depends on how you ask them," *Judgment and Decision Making*, 2012, *7* (6), 741–745.

Facchini, Giovanni, Yotam Margalit, and Hiroyuki Nakata, "Countering Public Opposition to Immigration: The Impact of Information Campaigns," 2016.

Faia, Ester, Andreas Fuster, Vincenzo Pezone, and Basit Zafar, "Biases in information selection and processing: Survey evidence from the pandemic," Technical Report, National Bureau of Economic Research 2021.

Falk, Armin and Florian Zimmermann, "A Taste for Consistency and Survey Response Behavior," *CESifo Economic Studies*, 2012, *59* (1), 181–193.

Fehr, Dietmar, Daniel Muller, Marcel Preuss et al., "Social Mobility Perceptions and Inequality Acceptance," *Working Paper*, 2021.

_ , Johanna Mollerstrom, and Ricardo Perez-Truglia, "Your Place in the World: The Demand for National and Global Redistribution," Technical Report, National Bureau of Economic Research 2019.

Ferrara, Eliana La, Alberto Chong, and Suzanne Duryea, "Soap operas and fertility: Evidence from Brazil," *American Economic Journal: Applied Economics*, 2012, *4* (4), 1–31.

Fetzer, Thiemo, Lukas Hensel, Johannes Hermle, and Christopher Roth, "Coronavirus perceptions and economic anxiety," *Review of Economics and Statistics*, 2020, pp. 1–36.

Finkelstein, Amy and Matthew J Notowidigdo, "Take-up and Targeting: Experimental Evidence from SNAP," *Quarterly Journal of Economics*, 2019, *134* (3), 1505–1556.

Fisman, Raymond, Keith Gladstone, Ilyana Kuziemko, and Suresh Naidu, "Do Americans want to tax capital? Evidence from online surveys," Technical Report, National Bureau of Economic Research 2017.

Fitzsimons, Emla, Bansi Malde, Alice Mesnard, and Marcos Vera-Hernandez, "Nutrition, information and household behavior: Experimental evidence from Malawi," *Journal of Development Economics*, 2016, *122*, 113–126.

Franklin, Simon, "Location, search costs and youth unemployment: experimental evidence from transport subsidies," *The Economic Journal*, 2017, *128* (614), 2353–2379.

Fuster, Andreas, Ricardo Perez-Truglia, Mirko Wiederholt, and Basit Zafar, "Expectations with Endogenous Information Acquisition: An Experimental Investigation," *Review of Economics and Statistics*, 10 2020, pp. 1–54.

Gaines, Brian J, James H Kuklinski, Paul J Quirk, Buddy Peyton, and Jay Verkuilen, "Same Facts, Different Interpretations: Partisan Motivation and Opinion on Iraq," *Journal of Politics*, 2007, *69* (4), 957–974.

Galashin, Misha, Martin Kanz, and Ricardo Perez-Truglia, "Macroeconomic Expectations and Credit Card Spending," Technical Report, National Bureau of Economic Research 2020.

Gärtner, Manja, Johanna Mollerstrom, and David Seim, "Income Mobility, Luck/Effort Beliefs, and the Demand for Redistribution: Perceptions and Reality," 2019.

Gennaioli, Nicola and Andrei Shleifer, "What comes to mind," *Quarterly Journal of Economics*, 2010, *125* (4), 1399–1433.

Gentzkow, Matthew and Jesse M Shapiro, "Media bias and reputation," *Journal of political Economy*, 2006, *114* (2), 280–316.

Gerber, Alan, Mitchell Hoffman, John Morgan, and Collin Raymond, "One in a million: Field experiments on perceived closeness of the election and voter turnout," *American Economic Journal: Applied Economics*, 2020, *12* (3), 287–325.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus, "Five Facts about Beliefs and Portfolios," *American Economic Review*, 2021, *111* (5), 1481–1522.

Gillen, Ben, Erik Snowberg, and Leeat Yariv, "Experimenting with Measurement Error: Techniques with Applications to the Caltech Cohort Study," *Journal of Political Economy*, 2019, *127* (4), 1826–1863.

Gong, Erick, "HIV testing and risky sexual behaviour," *The Economic Journal*, 2015, *125* (582), 32–60.

Graham, Matthew H, "Self-awareness of Political Knowledge," *Political Behavior*, 2018, pp. 1–22.

Greaves, Ellen, Iftikhar Hussain, Birgitta Rabe, Imran Rasul et al., "Parental Responses to Information About School Quality: Evidence from Linked Survey and Administrative Data," Technical Report, Institute for Social and Economic Research 2019.

Grewenig, Elisabeth, Philipp Lergetporer, Katharina Werner, and Ludger Woessmann, "Incentives, search engines, and the elicitation of subjective beliefs: evidence from representative online survey experiments," *Journal of Econometrics*, 2021.

_ , _ , Lisa Simon, Katharina Werner, and Ludger Woessmann, "Can Online Surveys Represent the Entire Population?," 2018.

Grigorieff, Alexis, Christopher Roth, and Diego Ubfal, "Does Information Change Attitudes Toward Immigrants?," *Demography*, 2020, *57* (3), 1–27.

Guess, Andrew M, Michael Lerner, Benjamin Lyons, Jacob M Montgomery, Brendan Nyhan, Jason Reifler, and Neelanjan Sircar, "A Digital Media Literacy Intervention Increases Discernment Between Mainstream and False News in the United States and India," *Proceedings of the National Academy of Sciences*, 2020.

Haaland, Ingar and Christopher Roth, "Labor market concerns and support for immigration," *Journal of Public Economics*, 2020, *191*, 104256.

_ and _ , "Beliefs about Racial Discrimination and Support for Pro-Black Policies," *Review of Economics and Statistics*, 2021.

Hager, Anselm, Lukas Hensel, Johannes Hermle, and Christopher Roth, "Strategic Interdependence in Political Movements and Countermovements," 2019.

_ , _ , _ , and _ , "Political Activists as Free-Riders: Evidence from a Natural Field Experiment," 2020.

_ , _ , _ , and _ , "Does Party Competition Affect Political Activism?," *Journal of Politics*, 2021.

Hainmueller, Jens and Daniel J Hopkins, "The Hidden American Immigration Consensus: A Conjoint Analysis of Attitudes Toward Immigrants," *American Journal of Political Science*, 2014, *59* (3), 529–548.

— **and Michael J Hiscox**, "Attitudes Toward Highly Skilled and Low-skilled Immigration: Evidence from a Survey Experiment," *American Political Science Review*, 2010, *104* (01), 61–84.

—**, Dominik Hangartner, and Teppei Yamamoto**, "Validating vignette and conjoint survey experiments against real-world behavior," *Proceedings of the National Academy of Sciences*, 2015, *112* (8), 2395–2400.

Hanna, Rema, Sendhil Mullainathan, and Joshua Schwartzstein, "Learning through noticing: theory and experimental evidence in farming," *Quarterly Journal of Economics*, 2014.

Hanspal, Tobin, Annika Weber, and Johannes Wohlfart, "Exposure to the COVID-19 Stock Market Crash and Its Effect on Household Expectations," *Review of Economics and Statistics*, 2020, pp. 1–45.

Hjort, Jonas, Diana Moreira, Gautam Rao, and Juan Francisco Santini, "How Research Affects Policy: Experimental Evidence from 2,150 Brazilian Municipalities," Technical Report, National Bureau of Economic Research 2019.

Hoffman, Elizabeth, Kevin McCabe, Keith Shachat, and Vernon Smith, "Preferences, property rights, and anonymity in bargaining games," *Games and Economic behavior*, 1994, *7* (3), 346–380.

Hoffman, Mitchell, "How is information valued? Evidence from framed field experiments," *The Economic Journal*, 2016, *126* (595), 1884–1911.

— **and Stephen V Burks**, "Worker overconfidence: Field evidence and implications for employee turnover and firm profits," *Quantitative Economics*, 2020, *11* (1), 315–348.

Hopkins, Daniel J, John Sides, and Jack Citrin, "The Muted Consequences of Correct Information About Immigration," *Journal of Politics*, 2019, *81* (1), 315–320.

Horton, John J, David G Rand, and Richard J Zeckhauser, "The Online Laboratory: Conducting Experiments in a Real Labor Market," *Experimental Economics*, 2011, *14* (3), 399–425.

Hossain, Tanjim and Ryo Okui, "The Binarized Scoring Rule," *Review of Economic Studies*, 2013, *80* (3), 984–1001.

Hoy, Christopher and Franziska Mager, "Can information about inequality and social mobility change preferences for redistribution? Evidence from randomized controlled trials in 11 high and middle-income countries," *Evidence from Randomized Controlled Trials in*, 2018, *11.*

Humphries, John Eric, Christopher Neilson, and Gabriel Ulyssea, "Information Frictions and Access to the Paycheck Protection Program," *Journal of Public Economics*, 2020.

Jamison, Julian C, Dean Karlan, and Pia Raffler, "Mixed method evaluation of a passive mHealth sexual information texting service in Uganda," Technical Report, National Bureau of Economic Research 2013.

Jensen, Robert, "The (perceived) returns to education and the demand for schooling," *Quarterly Journal of Economics*, 2010, *125* (2), 515–548.

Jessoe, Katrina and David Rapson, "Knowledge is (less) power: Experimental evidence from residential energy use," *American Economic Review*, 2014, *104* (4), 1417–38.

Kane, John V and Jason Barabas, "No Harm in Checking: Using Factual Manipulation Checks to Assess Attentiveness in Experiments," *American Journal of Political Science*, 2019, *63* (1), 234–249.

Karadja, Mounir, Johanna Mollerstrom, and David Seim, "Richer (and Holier) Than Thou? The Effect of Relative Income Improvements on Demand for Redistribution," *Review of Economics and Statistics*, 2017, *99* (2), 201–212.

Kendall, Chad, Tommaso Nannicini, and Francesco Trebbi, "How Do Voters Respond to Information? Evidence from a Randomized Campaign," *American Economic Review*, 2015, *105* (1), 322–53.

Kerwin, Jason, "Scared straight or scared to death? The effect of risk beliefs on risky behaviors," *The Effect of Risk Beliefs on Risky Behaviors (February 9, 2018)*, 2018.

Kerwin, Jason T and Natalia Ordaz Reynoso, "You Know What I Know: Interviewer Knowledge Effects in Subjective Expectation Elicitation," *Demography*, 2021.

Khan, Adnan, Sanval Nasim, Mahvish Shaukat, and Andreas Stegmann, "Building Trust in the State with Information: Evidence from Urban Punjab," 2020.

Kiessling, Lukas, "How Do Parents Perceive the Returns to Parenting Styles and Neighborhoods?," *Working Paper*, 2021.

Korlyakova, Darya, "Learning about Ethnic Discrimination from Different Information Sources," CERGE-EI Working Papers wp689, The Center for Economic Research and Graduate Education - Economics Institute, Prague March 2021.

Kremer, Michael and Edward Miguel, "The illusion of sustainability," *Quarterly journal of economics*, 2007, *122* (3), 1007–1065.

_ , Gautam Rao, and Frank Schilbach, "Behavioral development economics," in "Handbook of Behavioral Economics: Applications and Foundations 1," Vol. 2, Elsevier, 2019, pp. 345–458.

Krosnick, Jon A, "Response strategies for coping with the cognitive demands of attitude measures in surveys," *Applied cognitive psychology*, 1991, *5* (3), 213–236.

Krupka, Erin L and Roberto A Weber, "Identifying Social Norms Using Coordination Games: Why Does Dictator Game Sharing Vary?," *Journal of the European Economic Association*, 2013, *11* (3), 495–524.

Kuziemko, Ilyana, Michael I Norton, Emmanuel Saez, and Stefanie Stantcheva, "How Elastic are Preferences for Redistribution? Evidence from Randomized Survey Experiments," *American Economic Review*, 2015, *105* (4), 1478–1508.

Laudenbach, Christine, Annika Weber, and Johannes Wohlfart, "Beliefs About the Stock Market and Investment Choices: Evidence from a Field Experiment," *Working Paper*, 2021.

Leiser, David and Shelly Drori, "Naïve understanding of inflation," *Journal of Socio-Economics*, 2005, *34* (2), 179–198.

Lergetporer, Philipp, Guido Schwerdt, Katharina Werner, Martin R West, and Ludger Woessmann, "How Information Affects Support for Education Spending: Evidence from Survey Experiments in Germany and the United States," *Journal of Public Economics*, 2018, *167*, 138–157.

_ , Katharina Werner, and Ludger Woessmann, "Does ignorance of economic returns and costs explain the educational aspiration gap? Evidence from representative survey experiments," *Economica*, 2018.

_ , _ , and _ , "Educational inequality and public policy preferences: Evidence from representative survey experiments," *Journal of Public Economics*, 2020, *188*, 104226.

_ , Ludger Woessmann et al., "The political economy of university tuition fees: Information provision and income contingency in representative survey experiments," in "Annual Conference 2016 (Augsburg): Demographic Change" number 145703 Verein für Socialpolitik/German Economic Association 2016.

_ , Marc Piopiunik, Lisa Simon et al., "Do Natives' Beliefs About Refugees' Education Level Affect Attitudes Toward Refugees? Evidence from Randomized Survey Experiments," Technical Report, CESifo Group Munich 2017.

Levine, Ruth and Molly Kinder, *Millions saved: proven successes in global health*, Vol. 3, Peterson Institute, 2004.

Link, Sebastian, Andreas Peichl, Christopher Roth, and Johannes Wohlfart, "Information Frictions among Firms and Households," *Working Paper*, 2021.

Madajewicz, Malgosia, Alexander Pfaff, Alexander Van Geen, Joseph Graziano, Iftikhar Hussein, Hasina Momotaj, Roksana Sylvi, and Habibul Ahsan, "Can information alone change behavior? Response to arsenic contamination of groundwater in Bangladesh," *Journal of development Economics*, 2007, *84* (2), 731–754.

Manski, Charles F, "Measuring Expectations," *Econometrica*, 2004, *72* (5), 1329–1376.

_ , "Survey measurement of probabilistic macroeconomic expectations: progress and promise," *NBER Macroeconomics Annual*, 2018, *32* (1), 411–471.

Martinez-Bravo, Monica and Andreas Stegmann, "In Vaccines We Trust? The Effects of Anti-Vaccine Propaganda on Immunisation: Evidence from Pakistan," *CEMFI*, 2017.

Miller, Grant, Áureo De Paula, and Christine Valente, "Subjective Expectations and Demand for Contraception," Technical Report, National Bureau of Economic Research 2020.

Mobius, Markus M, Muriel Niederle, Paul Niehaus, and Tanya S Rosenblat, "Managing self-confidence: Theory and experimental evidence," Technical Report, National Bureau of Economic Research 2015.

Mummolo, Jonathan and Erik Peterson, "Demand Effects in Survey Experiments: An Empirical Assessment," *American Political Science Review*, 2019, *113* (2), 517–529.

Nyhan, Brendan, "Facts and Myths about Misperceptions," *Journal of Economic Perspectives*, 2020, *34* (3).

_ , "Why the backfire effect does not explain the durability of political misperceptions," *Proceedings of the National Academy of Sciences*, 2021, *118* (15).

_ **and Jason Reifler**, "When Corrections Fail: The Persistence of Political Misperceptions," *Political Behavior*, 2010, *42*, 303–300.

— **and** — , "Does correcting myths about the flu vaccine work? An experimental evaluation of the effects of corrective information," *Vaccine*, 2015, *33* (3), 459–464.

— **,** — **, Sean Richey, and Gary L Freed**, "Effective messages in vaccine promotion: a randomized trial," *Pediatrics*, 2014, *133* (4), e835–e842.

Orkin, Kate, "Everybody Loves a Winner: A Field Experiment Providing Information on Polls in South Africa," *Working Paper*, 2019.

Palfrey, Thomas R and Stephanie W Wang, "On eliciting beliefs in strategic games," *Journal of Economic Behavior & Organization*, 2009, *71* (2), 98–109.

Perez-Truglia, Ricardo and Guillermo Cruces, "Partisan Interactions: Evidence from a Field Experiment in the United States," *Journal of Political Economy*, 2017, *125* (4), 1208–1243.

— **and Ugo Troiano**, "Shaming tax delinquents," *Journal of Public Economics*, 2018, *167*, 120–137.

Peterson, Erik and Shanto Iyengar, "Partisan Gaps in Political Information and Information-Seeking Behavior: Motivated Reasoning or Cheerleading?," *American Journal of Political Science*, 2020.

Prina, Silvia and Heather Royer, "The importance of parental knowledge: Evidence from weight report cards in Mexico," *Journal of Health Economics*, 2014, *37*, 232–247.

Prior, Markus, Gaurav Sood, Kabir Khanna et al., "You Cannot be Serious: The Impact of Accuracy Incentives on Partisan Bias in Reports of Economic Perceptions," *Quarterly Journal of Political Science*, 2015, *10* (4), 489–518.

Qian, Wei, "House Price Expectations and Consumption - A Survey-based Experiment," *Working Paper*, 2019.

Rafkin, Charlie, Advik Shreekumar, and Pierre-Luc Vautrey, "When Guidance Changes: Government Inconsistency andPublic Beliefs," *Journal of Public Economics*, 2021, *191*, 104256.

Riley, Emma, “Role models in movies: The impact of Queen of Katwe on students’ educational attainment,” *Working Paper*, 2017.

Roth, Christopher and Johannes Wohlfart, “How do expectations about the macroeconomy affect personal expectations and behavior?,” *Review of Economics and Statistics*, 2020, *102* (4), 731–748.

_ , Sonja Settele, and Johannes Wohlfart, “Beliefs about public debt and the demand for government spending,” *Journal of Econometrics*, 2021.

_ , _ , and _ , “Risk Exposure and Acquisition of Macroeconomic Information,” *American Economic Review: Insights*, 2021.

Schwardmann, Peter and Joel Van der Weele, “Deception and Self-deception,” *Nature human behaviour*, 2019, *3* (10), 1055–1061.

Settele, Sonja, “How Do Beliefs About the Gender Wage Gap Affect the Demand for Public Policy?,” *Available at SSRN 3382325*, 2020.

_ and Cortnie Shupe, “Lives or Livelihoods? Perceived Tradeoffs and Public Demand for Non-Pharmaceutical Interventions,” *Perceived Tradeoffs and Public Demand for Non-Pharmaceutical Interventions (April 30, 2020)*, 2020.

Shacham, Enbal, Susan Godlonton, and Rebecca L Thornton, “Perceptions of male circumcision among married couples in rural Malawi,” *Journal of the International Association of Providers of AIDS Care (JIAPAC)*, 2014, *13* (5), 443–449.

Shrestha, Maheshwor, “Get Rich or Die Tryin’: Perceived Earnings, Perceived Mortality Rates, and Migration Decisions of Potential Work Migrants from Nepal,” *The World Bank Economic Review*, 2020, *34* (1), 1–27.

Snyder, Mark, “Self-monitoring of expressive behavior,” *Journal of personality and social psychology*, 1974, *30* (4), 526.

Stantcheva, Stefanie, “Understanding Economic Policies: What do People Know and How Can they Learn?,” Technical Report, Harvard University Working Paper 2020.

Thaler, Michael, "The "Fake News" Effect: An Experiment on Motivated Reasoning and Trust in News," Technical Report, Working Paper 2021.

Trautmann, Stefan T. and Gijs van de Kuilen, "Belief Elicitation: A Horse Race among Truth Serums," *The Economic Journal*, 9 2014, *125* (589), 2116–2135.

Tversky, Amos and Daniel Kahneman, "Judgment Under Uncertainty: Heuristics and Biases," *Science*, 1974, *185* (4157), 1124–1131.

Wang, Stephanie W, "Incentive effects: The case of belief elicitation from individuals in groups," *Economics Letters*, 2011, *111* (1), 30–33.

Wiswall, Matthew and Basit Zafar, "Determinants of college major choice: Identification using an information experiment," *Review of Economic Studies*, 2014, *82* (2), 791–824.

_ **and** _ , "How Do College Students Respond to Public Information About Earnings?," *Journal of Human Capital*, 2015, *9* (2), 117–169.

_ **and** _ , "Human capital investments and expectations about career and family," Technical Report, National Bureau of Economic Research 2016.

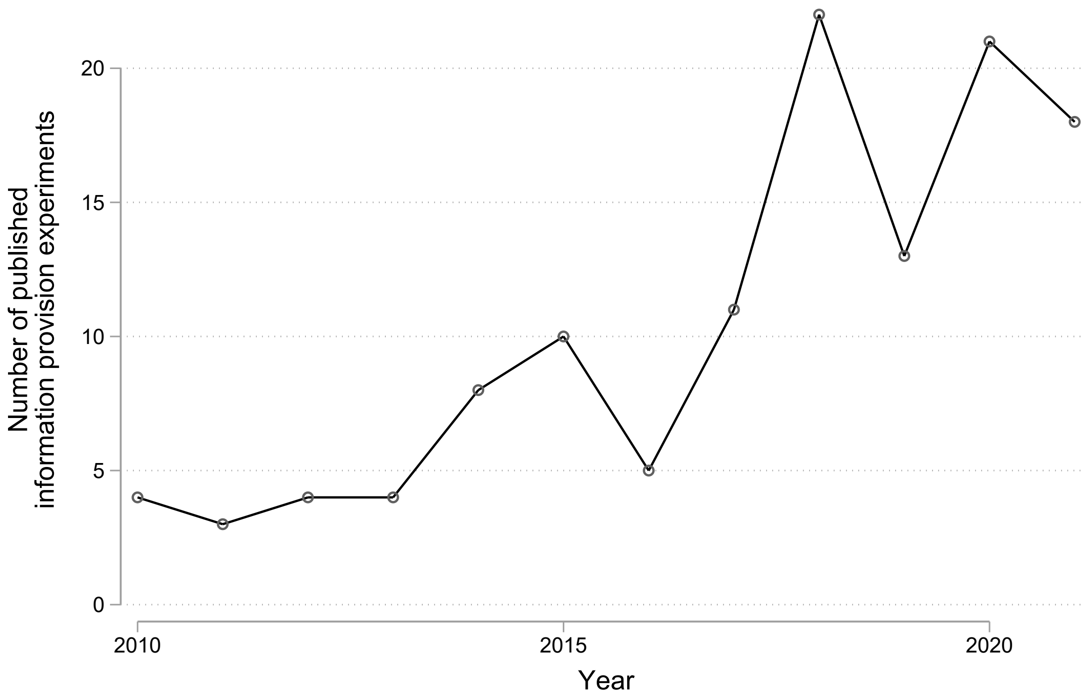
_ **and** _ , "Preference for the workplace, investment in human capital, and gender," *Quarterly Journal of Economics*, 2017, *133* (1), 457–507.

Yanagizawa-Drott, David, "Propaganda and conflict: Evidence from the Rwandan genocide," *Quarterly Journal of Economics*, 2014, *129* (4), 1947–1994.

Zimmermann, Florian, "The Dynamics of Motivated Beliefs," *American Economic Review*, February 2020, *110* (2), 337–61.

Zizzo, Daniel John, "Experimenter Demand Effects in Economic Experiments," *Experimental Economics*, 2010, *13* (1), 75–98.

Figure 1: Number of information provision experiments published in leading journals since 2010



Notes: This figure shows the number of information provision experiments published in leading journals since 2010. For 2021, publications and forthcoming papers as of mid-April are included. The figure is based on publications in the following journals: American Economic Review, American Economic Journal: Applied Economics, American Economic Journal: Economic Policy, American Economic Journal: Macroeconomics, Econometrica, Economic Journal, Journal of Development Economics, Journal of Political Economy, Journal of Public Economics, Journal of the European Economic Association, Review of Economics and Statistics, and the Review of Economic Studies. To identify articles, we used Google Scholar to search for all articles published in these journals since 2010 containing the words information, treatment, beliefs or expectations, and experiment, and then verified which of the search results featured an information provision experiment. We also supplemented with papers covered in our review that were not captured using this search algorithm. This figure does not include information provision experiments operating in a laboratory setting in which respondents receive information about features of the laboratory environment or the behavior of other participants in the lab. The figure also does not include information provision experiments in which respondents are informed about product characteristics.

Table 1: Estimated learning rates from quantitative information

Authors (year), journal, exhibit	Context	Type of information	Immediate LR	Time lapse	Follow-up LR	Persistence
Armantier et al. (2016), REStat, Table 2	Inflation expectations	Average professional forecast (SPF)	0.39	No follow-up	NA	NA
Armona et al. (2019), REStud, Table 9	House price expectations	1-year and 5-year past house price growth	0.18 (returns) 0 (5-year)	2 months	0.13	72%
Bleemer and Zafar (2018), JPubEc, Table 7	Perceived returns / cost college education own child	Population-level returns/cost	0.18 (returns) 0.35 (cost)	2 months	No learning rate for follow-up	NA
Bottan and Perez-Truglia (2020), REStat, Figure 2	Perceived city-level earnings rank and costs of living	Tailored based on potential own income and statistics from ACS or CPS	Earnings rank: 0.87 Cost of living: 0.88	5 weeks	Earnings rank: 0.63 Cost of living: 0.75	76% 85%
Cavallo et al. (2017), AEJ: Macro, Table 1	Inflation expectations	Statistics such as the current inflation rate or price changes of selected products	Statistics: 0.84 /0.43 (US/Arg.) Prices: 0.69 /0.46 (US/Arg.)	2 months/ 4 months (US/ Argentina)	Statistics: 0.36 /NA (US/Arg.) Prices: 0.34/ 0.21 (US/Arg.)	43% 49% 46%
Coibion et al. (2020e), QJE, Table 3	Firms' first and higher-order inflation expectations	Avg. first- and higher-order exp. of other firms	1st-ord. exp.: 0.31/ 0.88 (others' 1st-/ higher-ord. exp.) higher-ord. exp.: 0.38/ 0.84 (others' 1st-/ higher-ord. exp.)	3 months	1st-ord. exp.: 0.39/ 0.82 (others' 1st-/ higher-ord. exp.) higher-ord. exp.: 0.28/ 0.79 (others' 1st-/ higher-ord. exp.)	126% 93% 74% 94%
Coibion and Gorodnichenko (2015, AER, Tables 6 and 7	Firms' expectations about inflation unemployment and GDP growth	Infl.: Prof. forecast or central bank target; Unempl. and GDP: Past 12 months	Infl: 0.66 (pooled) Unempl.: 0.35 GDP: 0.44	6 months	0	0%
Fuster et al. (2020), REStat, Figures 3 and A.2	House price expectations	Choice btw. expert forecast, past 1 year and past 10 year house price growth	0.38 (based on preferred source of information)	4 months	0.17	45%
Roth et al. (2021a), JEconometrics, Table A.18	Beliefs about the debt-to-GDP ratio	Factual information	0.62	4 weeks	0.21	34%
Roth et al. (2021b), AER: Insights, Table 2	Perceived unemployment risk next recession	Change in unempl. rate in own demographic group	0.49	No follow-up	NA	NA
Roth and Wohlfart (2020), REStat, Table 1	Recession expectations	Individual professional forecasts (SPF)	0.32	2 weeks	0.13	41%
Wiswall and Zafar (2014), REStud, Table 4	Own earnings prospects by college major	Population-level earnings by college major	0.08	2 years	No learning rate for follow-up	NA
Wiswall and Zafar (2015), JHumanCapital, Table 7	Own earnings prospects by college major	Population-level earnings by college major	0.34	No follow-up	NA	NA

Notes: This table provides an overview of estimated learning rates in published papers that study how individuals update their beliefs in response to the provision of quantitative information and that calculate learning rates. The table shows both the immediate learning rate measured in the main survey, and the medium-run learning rate as measured with posterior beliefs elicited in a follow-up survey. For Coibion et al. (2020e), the learning rate is calculated as one minus the weight put on the prior, scaled by the weight put on the prior in the control group. For Coibion and Gorodnichenko (2015), the learning rate is calculated as one minus the weight put on the prior.